

---

# **DIPLOMARBEIT**

---

Herr  
**Stefan Bernstein**

**Entwicklung einer grafischen  
Oberfläche zum Browsen und  
Labeln von Trainingsbildern**

2010



---

# **DIPLOMARBEIT**

---

## **Entwicklung einer grafischen Oberfläche zum Browsen und Labeln von Trainingsbildern**

Autor:

**Stefan Bernstein**

Studiengang:

Multimediatechnik

Seminargruppe:

MK06W2

Erstprüfer:

Prof. Dr. rer. nat. Günter Werner

Zweitprüfer:

Dipl.-Inf. Julia Möhrmann

Mittweida, August 2010



---

## **Bibliografische Angaben**

Bernstein, Stefan: Entwicklung einer grafischen Oberfläche zum Browsen und Labeln von Trainingsbildern, 45 Seiten, 31 Abbildungen, Hochschule Mittweida (FH), Fakultät Elektro- und Informationstechnik

Diplomarbeit, 2010

Dieses Werk ist urheberrechtlich geschützt.

## **Referat**

Die Diplomarbeit beschäftigt sich mit der Erstellung einer Anwendung um große Bilddatensätze für Bilderkennungssysteme zu kategorisieren. Die Bilder sollen mithilfe einer selbstorganisierenden Karte („self organizing map“) vorsortiert angeordnet werden. Es sind Funktionen zu entwickeln, welche den Nutzer beim labeln unterstützen. Ziel ist es, mit der Anwendung Bilder schneller und einfacher labeln zu können, als per Hand bzw. ohne Vorsortierung. Mit einer anschließenden Benutzerstudie wird die Erreichung dieses Ziels überprüft.



# I. Inhaltsverzeichnis

Inhaltsverzeichnis .....	I
Abbildungsverzeichnis .....	II
Tabellenverzeichnis .....	III
Abkürzungsverzeichnis .....	IV
Vorwort .....	V
1 Einleitung .....	1
2 Problem und Aufgabe .....	3
2.1 Problem .....	3
2.2 Aufgabenstellung .....	4
2.3 Sortierungsverfahren .....	4
2.4 Darstellung .....	5
2.5 Zusammenfassung der Anforderungen .....	6
3 Implementierung .....	7
3.1 Vorarbeit .....	7
3.2 Aufbau der GUI .....	7
3.3 SOM-Darstellung .....	9
3.4 Anordnung der Bilder .....	12
3.5 Maus- und Tastaturinteraktion .....	13
3.6 Visuelle Hilfen .....	15
4 Evaluation .....	17
4.1 Testaufbau und -ablauf .....	17
4.2 Testauswertung .....	21
4.3 Anpassungen nach der Evaluation .....	39
5 Zusammenfassung und Ausblick .....	41
Literaturverzeichnis .....	43





## II. Abbildungsverzeichnis

2.1	SOM Berechnungsbeispiel - Suche nach ähnlichstem Gewichtsvektor .....	4
2.2	SOM Berechnungsbeispiel - Anpassung des Gewichtsvektors .....	5
3.1	Inteface.....	8
3.2	SOM Anordnungsbeispiel - Kanten .....	10
3.3	SOM Standardanordnung .....	10
3.4	SOM verschobene Anordnung .....	11
3.5	U-Matrix Farbverläufe .....	12
3.6	Bilder in Rasteranordnung .....	13
3.7	Bilder in Kreisanordnung .....	13
3.8	Auswahlrahmen .....	14
3.9	Bild im Status „Aktiv“ .....	16
3.10	Bild im Status „Inaktiv“ .....	16
4.1	Beispiele COIL Datensatz.....	18
4.2	Beispiele Fensterdatensatz .....	18
4.3	Beispiele Parkplatzdatensatz .....	19
4.4	Diagramm - Zeit mit und ohne Vorsortierung.....	22
4.5	Diagramm - Benotung der Geschwindigkeit mit und ohne Vorsortierung .....	23
4.6	Diagramm - Benotung der Schwierigkeit mit und ohne Vorsortierung .....	24
4.7	Diagramm - Richtigkeit der vergebenen Labels mit und ohne Vorsortierung .....	25
4.8	Schwer zu findendes, falsch zugeordnetes Bilder eines Knotens.....	26
4.9	Diagramm - Zeit SOM und SOM mit verschobenen Knoten.....	28
4.10	Diagramm - Fehler SOM und SOM mit verschobenen Knoten .....	29
4.11	Diagramm - Beste Visualisierung Parkplatzdatensatz .....	30
4.12	Diagramm - Zeit SOM und SOM mit farbigen Hexagons .....	31
4.13	Diagramm - Fehler Standard SOM und SOM mit farbigen Hexagons .....	32
4.14	Diagramm - Beste Visualisierung COIL-100- und Fensterdatensatz.....	33
4.15	Diagramm - Zeit mit und ohne Bildstatusfunktion .....	35

4.16 Diagramm - Bewertung von Geschwindigkeit und Schwierigkeit mit und ohne Bildstatusfunktion .....	36
4.17 Diagramm - Bewertung der Kompassfunktion .....	37
4.18 Diagramm - Vergleich Anzahl der Zoom- und Verschiebevorgänge.....	38
4.19 Diagramm - Bestes Auswahlwerkzeug .....	39

---

## III. Tabellenverzeichnis

4.1 Tests pro Datensatz .....	19
4.2 Vorhandene Testwerte pro Darstellung .....	21
4.3 Verbesserungsvorschläge .....	40



## IV. Abkürzungsverzeichnis

BMU .....	Best Matching Unit, Seite 4
Coalda .....	Coreference Annotation of Large Datasets, Seite 7
HSB .....	Hue Saturation Brightness, Seite 12
SOM .....	Self Organizing Map, Seite 4
U-Matrix .....	Unified Distance Matrix, Seite 5
ZUI .....	zoomable user interface bzw. zooming user interface, Seite 6



## **V. Vorwort**

Vielen Dank an meine Familie, welche mich während der Arbeit unterstützt hat. Ich danke meiner Betreuerin Julia Möhrmann für die vielen Tipps und Denkanstöße. Weiterhin danke ich Andre Burkovski für den SOM-Server und Stefanie Wiltrud Kessler für den hervorragend kommentierten Code. Vielen Dank allen Teilnehmern der Benutzerstudie für ihr Durchhaltevermögen. Außerdem Danke an die Korrekturleser: Carsten Münch, Christian Bernstein, Christian Kessler und Cong Wang.





# 1 Einleitung

Das Erstellen von Bilderkennungssystemen ist sehr zeit- und arbeitsaufwändig. Von den zu erkennenden Objekten wird eine große Anzahl Trainingsbilder benötigt. Jeder zu erkennende Zustand der Objekte (z.B. Fenster offen, geschlossen) muss bei verschiedenen Beleuchtungsbedingungen (z.B. morgens, mittags, abends, künstliche Beleuchtung) abgebildet werden. Daraufhin muss dieser, oft bis zu 1000 Bilder beinhaltende, Datensatz in Kategorien unterteilt werden. Diese spiegeln die verschiedenen Zustände der Objekte wieder und werden für das Training des Bilderkennungssystems benötigt. In dieser Arbeit soll untersucht werden, inwiefern die Vorsortierung eines solchen Bilddatensatzes das Kategorisieren für den Nutzer vereinfacht. Es sollen verschiedene Möglichkeiten entwickelt werden einen, mit einer selbstorganisierenden Karte („self organizing map“, „som“) sortierten Datensatz darzustellen. Dabei geht es vor allem darum dem Nutzer ein schnelleres und leichteres Kategorisieren seiner Bilder zu ermöglichen. Grundlage für die Arbeit sind [Heidemann et al., 2003] und [Bekel et al., 2005].

Neben der Darstellung der vorsortierten Bilder, geht es um die Entwicklung einer einfach zu nutzenden Benutzeroberfläche. Mit Hilfe einer Benutzerstudie soll anschließend die vorsortierte Darstellung mit der unsortierten Darstellung verglichen werden.



## 2 Problem und Aufgabe

### 2.1 Problem

Die Arbeit mit Bilderkennungssystemen wird durch den großen Aufwand bei der Erstellung erschwert. Sowohl die Zeit, als auch die manuelle Arbeit, welche nötig ist um Bilddatensätze für das System vorzubereiten, muss reduziert werden. Vor allem das Kategorisieren der einzelnen Bilder stellt ein Problem dar, weil es viel Arbeit auf Seiten des Benutzers erfordert. Das Sammeln der Bilder dagegen, kann größtenteils unbeaufsichtigt erledigt werden. Da das Kategorisieren gerade dazu dient ein Bilderkennungssystem zu entwickeln, kann es nicht automatisiert werden. Manuelle Arbeit durch den Benutzer ist nötig. Diese sollte jedoch einfacher und schneller zu erledigen sein.

Meistens ist die Anzahl der verschiedenen Bildkategorien sehr gering. Zum Beispiel soll ein System zur Überwachung einer Tür nur zwischen den Zuständen „offen“ und „geschlossen“ unterscheiden. Auf den Bildern lassen sich diese beiden Zustände auch sehr einfach erkennen. Trotzdem muss der Benutzer jedes Bild einzeln überprüfen und kategorisieren. Das kommt daher, dass die Bilder im Dateimanager nicht nach Inhalt, sondern nur nach Dateinamen oder Erstellungszeit, sortiert werden können. Durch die Sortierung nach Erstellungszeiten könnte eine gewisse Gruppierung und damit Vereinfachung erreicht werden. Dies funktioniert allerdings nur, wenn die Kategorien in Zusammenhang mit der Aufnahmezeit stehen. Zum Beispiel füllt sich ein Parkplatz meist nur langsam. Darum werden zeitlich aufeinander folgende Bilder oft ähnlich sein. Außerdem ist ein Firmenparkplatz meistens nur während der Arbeitszeiten belegt. Somit ließen sich die Kategorien „voll“ und „leer“ grob Anhand der Tageszeiten zuordnen. Bei dem Türbeispiel funktioniert dieser Ansatz aber schon nicht mehr, da die Zustandsänderungen sehr schnell erfolgen. Die Sortierung nach Erstellungszeit stellt darum keine ausreichende Lösung für das Problem dar.

Beim Umgang mit sehr vielen Bildern besteht außerdem das Problem der Unübersichtlichkeit. Es ist schwierig alle Bilder auf einmal, in sinnvoller Größe anzuzeigen. Jedoch ist beim Kategorisieren wichtig, dass der Benutzer einen Überblick über alle verfügbaren Bilder erhält. Denn sonst kann er nicht entscheiden, ob zwei Bilder unterschiedlich genug sind um in verschiedene Kategorien zu gehören oder nicht. Die Anordnung in Listen oder Rastern, wie sie oft in Dateimanagern verwendet wird, ist dabei nicht sinnvoll. Die Dauer zum Scrollen von Anfang bis Ende der Darstellung ist viel zu lang, um sich dabei alle Arten von Bildern zu merken.

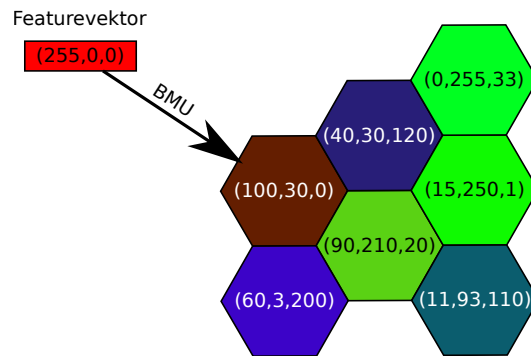


Abbildung 2.1: Beispielhafter Ausschnitt einer SOM Berechnung mit RGB-Werten als Featurevektoren. Die Knotenpunkte (rechts) besitzen Gewichtsvektoren mit zufällig initialisierten Werten. Der dem Featurevektor (Rot) ähnlichste Gewichtsvektor wird gesucht. Die Farben der Hexagons entsprechen den Gewichtsvektorwerten der Knoten.

## 2.2 Aufgabenstellung

Es soll eine Anwendung entwickelt werden, welche Bilder nach Ähnlichkeit ihrer Inhalte sortiert und anordnet. Zur Sortierung wird eine selbstorganisierende Karte verwendet. Die sortierten Bilder sollen auf übersichtliche Art dargestellt werden. Es sind Funktionen zu entwickeln, welche den Benutzer beim Kategorisieren der Bilder unterstützen und damit die Arbeit schneller und einfacher machen.

## 2.3 Sortierungsverfahren

Um die Bilder anhand von inhaltlichen Ähnlichkeiten zu sortieren, soll eine selbstorganisierende Karte („self organizing map“, „SOM“) verwendet werden.

Selbstorganisierende Karten bilden hochdimensionale Eingangsdaten („Featurevektoren“) als niedrigdimensionale (meist 2D) Rasterdarstellung ab. Dabei können Beziehungen zwischen einzelnen Featurevektoren und Gruppen innerhalb der Eingangsdaten leicht erkannt werden. [Kohonen, 1990]

Eine SOM besteht aus Knotenpunkten („Knoten“, „nodes“), welche eine Position in einem Raster und einen Gewichtsvektor besitzen. Der Gewichtsvektor hat die selbe Dimension wie die Featurevektoren und wird meist mit zufälligen Werten initialisiert. Beim Training der selbstorganisierenden Karte wird die euklidische Distanz zwischen jedem Featurevektor und jedem Gewichtsvektor berechnet. Je kürzer diese ist, desto ähnlicher ist der Featurevektor dem Gewichtsvektor. Der Knotenpunkt, dessen Gewichtsvektor die kürzeste Distanz zum aktuellen Featurevektor hat, wird ausgewählt. Dieser Knoten wird dann als „best matching unit“ („bmu“) dieses Featurevektors bezeichnet. (Siehe Abbildung 2.1 auf Seite 4)

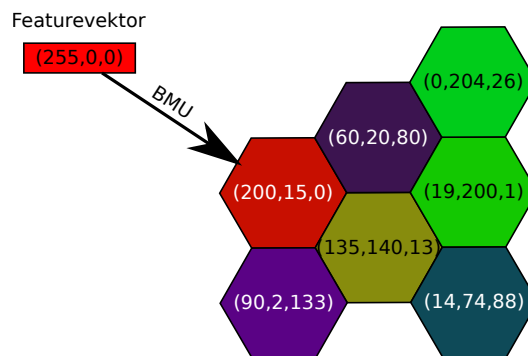


Abbildung 2.2: Beispielhafter Ausschnitt einer SOM Berechnung mit RGB-Werten als Featurevektoren. Der Gewichtsvektorwert des BMU-Knotenpunktes wird dem Featurevektor angenähert. Die Farben der Hexagons entsprechen den Gewichtsvektorwerten der Knoten. Durch den Vergleich mit Abbildung 2.1 wird der Annäherungsvorgang deutlich.

Wenn der ähnlichste Gewichtsvektor eines Featurevektors gefunden wurde, wird der Wert dieses Gewichtsvektors dem Wert des Featurevektors angenähert. Dasselbe geschieht in fallender Stärke mit den Gewichtsvektoren der umliegenden Knotenpunkte. (Siehe Abbildung 2.2 auf Seite 5)

Dieser Vorgang wird für alle übrigen Featurevektoren wiederholt. Anschließend sind untereinander ähnliche Featurevektoren den selben oder nah beieinander liegenden Knotenpunkten zugeordnet.

Zusätzlich zu den Gewichtsvektoren erhält jeder Knotenpunkt und jede Verbindung zwischen zwei Knotenpunkten („Kante“, „edge“) einen Wert, welcher die Ähnlichkeit darstellt. Dieser Wert wird U-Matrix Wert genannt. Bei Knotenpunkten gibt der U-Matrix Wert die mittlere Distanz des Gewichtsvektors dieses Knotens zu allen umliegenden Knoten an. Je ähnlicher der Knoten zu den umliegenden Knoten ist, desto geringer ist der Wert. Bei Kanten wird die Distanz zwischen den beiden verbundenen Knotenpunkten angegeben. Je ähnlicher die beiden Knoten sind, desto geringer ist der U-Matrix Wert der Kante.

Zur Berechnung der SOM wird eine bereits vorhandene Implementierung verwendet. Diese ist ein SOM-Server, welcher Featurevektoren aus einer Datenbank liest und die berechnete SOM in die Datenbank schreibt. Zur Erstellung des Featurevektors für jedes Bild steht ein Matlab Skript zur Verfügung, welches Histogramme als Features verwendet.

## 2.4 Darstellung

Um die Übersichtlichkeit zu erhöhen sollen die angeordneten Bilder dem Benutzer in Form einer zoombaren Oberfläche („ZUI“) präsentiert werden. Diese ermöglicht es die

Darstellung wie durch eine Kamera zu betrachten. Der Nutzer kann die Entfernung zu den Bildern vergrößern und verkleinern, um sich einen Überblick zu verschaffen. Der sichtbare Bereich der Darstellung kann verschoben werden. Durch die zoombare Oberfläche soll der Eindruck eines unendlich großen Darstellungsbereiches erzeugt werden. Außerdem kann der Benutzer die grobe Anordnung aller Bilder leichter im Gedächtnis behalten.

## 2.5 Zusammenfassung der Anforderungen

Die Anwendung muss folgende Funktionen enthalten:

- Laden von Featurevektoren eines Bilddatensatzes in eine Datenbank
- Starten der SOM Berechnung auf einem SOM Server
- Laden der Ergebnisse einer SOM Berechnung aus einer Datenbank
- Übersichtliche Darstellung und Anordnung von Bildern in Form einer ZUI
- Einfache Auswahl und Kategorisierung von Bildern
- Speichern der Kategorien in einer Datenbank

Die eigentliche Berechnung der SOM und der Featurevektoren ist nicht Teil dieser Arbeit.

## 3 Implementierung

### 3.1 Vorarbeit

Zur Einarbeitung in die Kommunikation mit dem SOM Server stand der kommentierte Quelltext eines anderen Projektes (Coalda [Kessler, 2010]) zur Verfügung. Aus diesem konnte auf die Struktur der SOM Daten in der Datenbank geschlossen werden. Einige Teile, wie die Klasse zum Verbindungsaufbau, wurden mit wenig Änderungen wiederverwendet. In einigen Punkten unterscheidet sich die Datenstrukturen des vorliegenden Projektes jedoch von der des Coalda Projektes. Darum wurden mehrere Klassen zum Laden der SOM Daten verändert oder neu implementiert.

Zu Beginn der Arbeit musste auch entschieden werden, ob zur Darstellung der SOM ein Framework verwendet werden soll. Im Coalda Projekt werden die SOM Daten mit dem Prefuse Framework gerendert [Heer J. et al., 2005]. Nach einigen Tests erschien Prefuse jedoch zu langsam bei der Darstellung. Denn neben der Knotenstruktur müssten auch noch ca. 800 Bilder gerendert werden. Außerdem bot eine Eigenentwicklung mehr Freiheiten für das Experimentieren mit neuen Funktionen.

### 3.2 Aufbau der GUI

Die Oberfläche (siehe Abbildung 3.1 auf Seite 8) besteht aus drei Bereichen: der SOM-Darstellung, der Labelleiste und der Kontrollleiste.

In der Kontrollleiste stehen die beiden Selektionswerkzeuge (Pfad und Rechteck) zur Auswahl. Diese werden in Abschnitt 3.5 auf Seite 13 genauer vorgestellt. Weiterhin kann die Zoomstufe der SOM-Darstellung mit dem Regler in der Mitte verändert werden. Der Regler auf der rechten Seite dient zum vergrößern und verkleinern der aktuellen Auswahl.

Im unteren Teil der Labelleiste werden alle verfügbaren Kategorien/Label angezeigt. Jedes Label besitzt einen Knopf, welcher mit dessen Farbe und Name beschriftet ist. Mit einem Klick auf diesen Knopf wird das Label den gerade ausgewählten Bildern zugewiesen. Falls die Bilder bereits ein Label besitzen, wird dieses durch das neue Label ersetzt. Jedes Bild kann nur zu einer Kategorie gehören/ein Label besitzen. Hinter jedem Labelknopf befindet sich ein weiterer Knopf um das Label komplett zu löschen. Dadurch wird das Label unwiderruflich aus der Liste der verfügbaren Labels gelöscht. War dieses Label bereits Bildern zugewiesen, so wird es von diesen entfernt. Um Label von Bildern zu entfernen ohne die Label zu löschen steht ein zusätzlicher Knopf zur Verfügung. Dieser befindet sich über den Labelknöpfen und heißt „clear label“, beziehungsweise

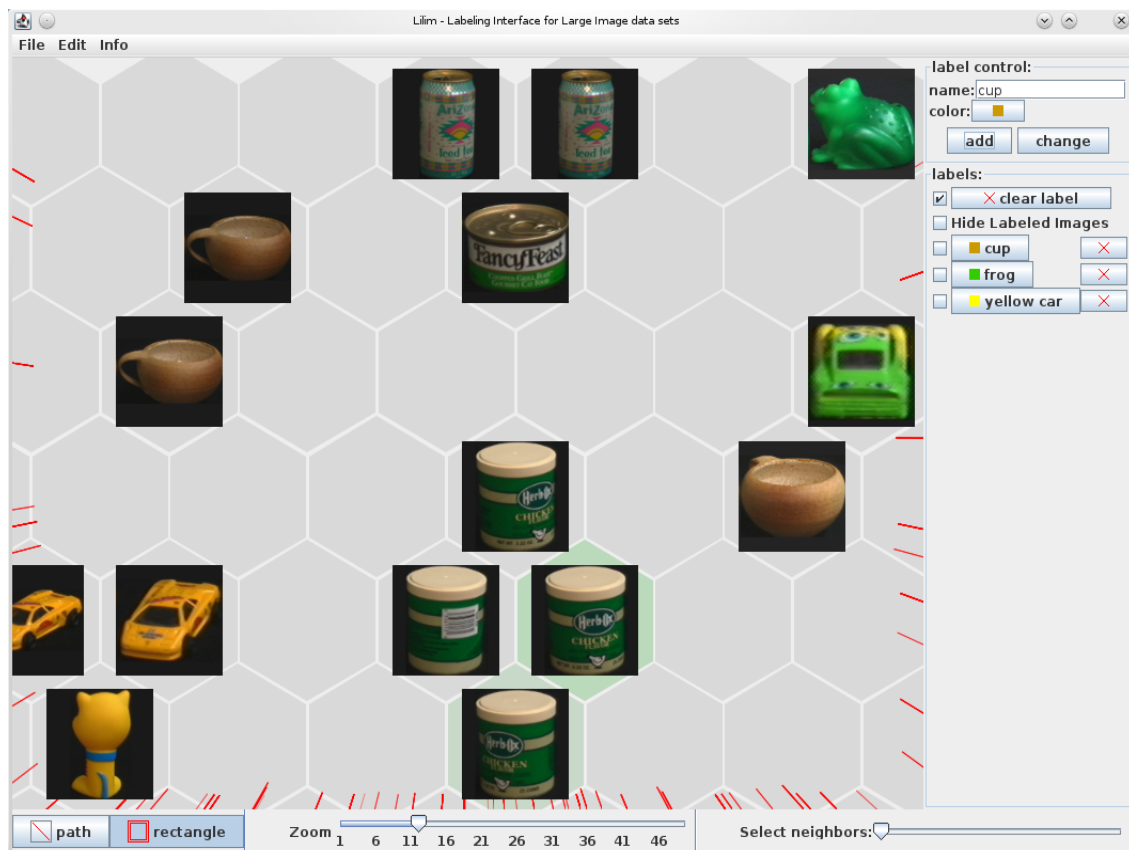


Abbildung 3.1: Oberfläche der Anwendung, bestehend aus SOM-Darstellung (links), Labelleiste (rechts), Kontrollleiste (unten) und Menüleiste (oben).



„Label löschen“ in der deutschen Version. Ein Klick auf diesen Knopf entfernt das jeweilige Label von allen gerade markierten Bildern und versetzt diese wieder in ungelabelten Zustand. Die Kontrollkästchen vor den Labelknöpfen dienen zum Statuswechsel aller Bilder mit diesem Label. Diese können entweder „Aktiv“ oder „Inaktiv“ sein. Die beiden Status werden im Abschnitt 3.6 auf Seite 15 näher erläutert. Über ein zusätzliches Kontrollkästchen besteht außerdem die Möglichkeit, alle gelabelten Bilder komplett unsichtbar zu machen.

Der obere Teil der Labelleiste dient zum Erstellen und Verändern der Labels. Der Benutzer kann einen Namen und eine Farbe wählen und damit ein neues Label erstellen. Allerdings darf kein Name doppelt verwendet werden. Bestehende Label werden bei einem Klick auf ihren Knopf in den Editiermodus versetzt. Der Name und die Farbe des angeklickten Labels erscheinen im oberen Teil der Labelleiste und können verändert werden. Durch einen Klick auf den „change“- beziehungsweise „ändern“-Knopf werden die Änderungen angewendet.

Die SOM-Darstellung füllt den verbleibenden Bereich des Anwendungsfensters. Zu sehen ist, je nach Zoomstufe, die komplette Karte oder nur ein Ausschnitt. Der Benutzer kann sich durch die Darstellung bewegen, Bilder selektieren, verschieben und labeln. (Siehe Abschnitt 3.5 auf Seite 13)

## 3.3 SOM-Darstellung

### 3.3.1 Standard

Die selbstorganisierende Karte besteht gewöhnlicher Weise aus einem quadratischen oder hexagonalen Raster. Dieses wird aus Knotenpunkten und den Verbindungen dazwischen gebildet (Siehe Abbildung 3.2 auf Seite 10). Jeder Knotenpunkt besitzt eine festgelegte horizontale und vertikale Position. Den Knotenpunkten wurden während der Berechnung Bilder, in Form von Featurevektoren, zugewiesen. Ähnliche Bilder wurden den selben Knoten zugeordnet. Ähnliche Knoten liegen außerdem meist in der selben Region. Dies ist jedoch nicht zwingend der Fall. Sich ähnelnde Knoten können auch in verschiedenen Regionen liegen (siehe Knoten mit Froschbildern in Abbildung 3.3). Diese Anordnung soll helfen ähnliche Bilder einfacher und schneller der selben Gruppe zuzuordnen.

Die Standard-SOM-Darstellung verwendet die vorgegebenen Positionen der Knoten und ordnet die zugehörigen Bilder entsprechend an. In Abbildung 3.3 auf Seite 10 ist eine solche Darstellung zu sehen. Jeder Schnittpunkt von Linien ist ein Knotenpunkt. Einigen wurden Bilder zugeordnet, anderen nicht. Da das Zoomlevel in der Abbildung gering ist, wird nur ein Bild pro Knoten angezeigt. Dieses wird Repräsentant genannt und in Abschnitt 3.4 auf Seite 12 näher erläutert.

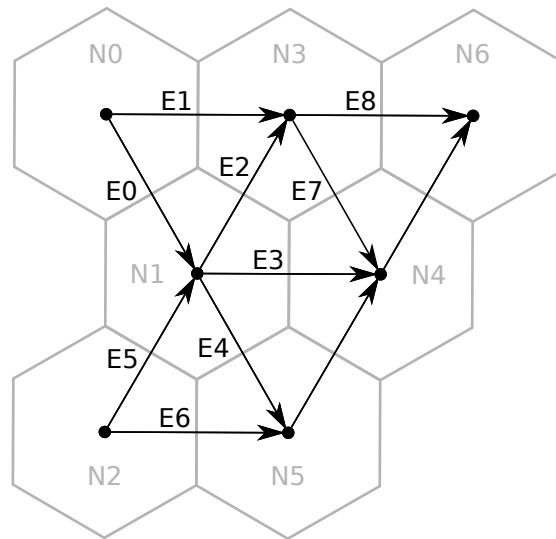


Abbildung 3.2: Beispielhafter Ausschnitt einer SOM. Die Knotenpunkte sind im hexagonalen Raster angeordnet. Gekennzeichnet werden sie durch den Punkt in der Mitte der Hexagons und einen Index mit vorstehendem N . Zwischen den Knoten befinden sich Kanten. Diese besitzen jeweils Start- und Endknoten, welche durch die Pfeilrichtung gekennzeichnet sind. Die Nummerierung der Kanten und Knoten ist identisch der Nummerierung der wirklichen SOM.



Abbildung 3.3: Beispiel für die normale SOM-Darstellung. Die Knoten sind im Standardraster angeordnet und ein Repräsentant der zugeordneten Bilder wird angezeigt. In der normalen Version werden die Kanten nicht angezeigt.



Abbildung 3.4: Beispiel für die verschobene SOM-Darstellung. Ausgehend vom Standardraster wurden die Knoten mit Hilfe der U-Matrix Werte der Kanten verschoben. Die Distanz zwischen zwei Knoten entspricht nun dem U-Matrix Wert der Kante. Für jeden Knoten ist nur ein Repräsentant aller zugehörigen Bilder sichtbar. In der normalen Version werden die Kanten nicht angezeigt.

### 3.3.2 Mit U-Matrix verschobene Darstellung

Zusätzlich zur normalen SOM-Darstellung können auch die U-Matrix Werte zur Anordnung der Knotenpunkte verwendet werden. Dazu wird die ursprüngliche Rasteranordnung als Ausgangspunkt verwendet. Bei dieser sind die Distanzen zwischen den Knoten gleich 1. Die Knoten werden anhand der U-Matrix Werte ihrer Kanten verschoben. Anschließend entsprechen die Distanzen der Knotenpunkte den U-Matrix Werten der Kanten. Da die U-Matrix Werte die Ähnlichkeit angeben, liegen ähnliche Knoten dichter bei einander als unterschiedliche Knoten. (Siehe Abbildung 3.4 auf Seite 11)

Diese Anordnung soll die Auswahl von zusammengehörigen Bildern weiter vereinfachen. Regionen mit ähnlichen Knoten werden deutlicher sichtbar als mit der Standard-SOM-Darstellung. Außerdem wird das Auswählen von Knoten einfacher.

### 3.3.3 U-Matrix Farbdarstellung

Alternativ zum Verschieben der Knotenpositionen können die Ähnlichkeiten zwischen Knotenpunkten durch Farben dargestellt werden. Dazu werden die U-Matrix Werte der Knotenpunkte verwendet um Farbwerte zu codieren. Die erzeugten Farben werden verwendet, um hinter jedem Knotenpunkt ein Hexagon einzufärben. Dieses soll helfen Gruppen ähnlicher Knotenpunkte schneller zu finden. Anfangs wurde ein Farbspektrum von Blau bis Rot gewählt, mit Blau für sehr ähnliche Knoten und Rot für sehr verschiedene Knoten. Dieses Spektrum wurde gewählt, da es häufig für die Visualisierung der U-Matrix verwendet wird. Da Rot aber auch für die Auswahlformen verwendet wurde, musste dieses Spektrum nachträglich verändert werden. Außerdem schienen die Farben zu

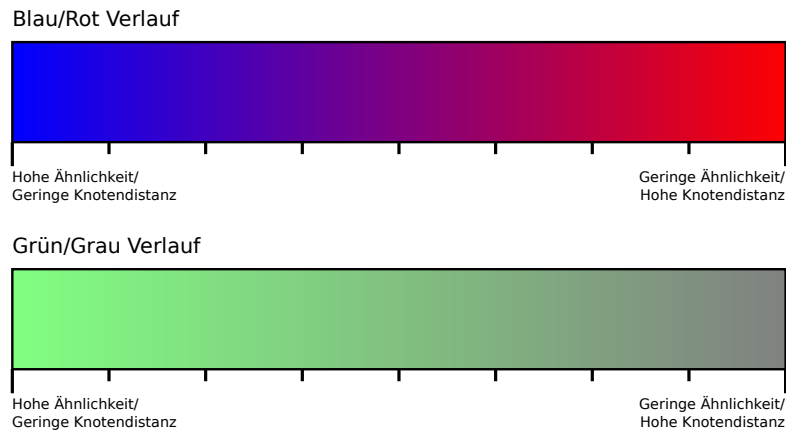


Abbildung 3.5: Farbverläufe zur Darstellung der Ähnlichkeiten/Distanzen zwischen Bildern. Oben der anfangs gewählte Blau/Rot Verlauf. Weil die Auswahlformen ebenfalls Rot verwenden, wurde der Grün/Grau Verlauf entworfen. Dieser soll außerdem weniger irritieren als der Blau/Rot Verlauf, da die resultierenden Farben unauffälliger sind.

auffällig, so dass der Benutzer irritiert wurde (siehe Abschnitt 4.2.2 auf Seite 27). Als Alternative wird darum ein Spektrum von Grün zu Grau verwendet. Dabei wird das HSB-Farbschema verwendet, um mit den U-Matrix Werten nur die Farbsättigung zu verändern. Daraus ergibt sich, dass die Sättigung des Grüntones bei ähnlichen Knotenpunkten höher wird und mit sinkender Ähnlichkeit geringer wird. Diese beiden Farbspektren sind dargestellt in Abbildung 3.5 auf Seite 12.

### 3.4 Anordnung der Bilder

Jedem Knotenpunkt kann eine beliebige Anzahl Bilder zugeordnet sein und kein Bild hat eine vordefinierte Position. Dies wirft das Problem auf, dass die Bilder um ihren Knoten angeordnet werden müssen ohne dabei unübersichtlich zu wirken. Aus diesem Grund wurde die Darstellung in Form einer zoombaren Oberfläche entwickelt. Die Anzahl der angezeigten Bilder sollte mit steigendem Zoomlevel wachsen, da nur noch ein kleiner Bereich der gesamten Darstellung sichtbar ist. Dazu wird von jedem Knoten das erste zugewiesene Bild als Repräsentant ausgewählt. Wenn der Zoomlevel einen bestimmten Wert unterschreitet, wird auf jedem Knoten nur dieser Repräsentant dargestellt. Wird der Wert überschritten, entfaltet sich der Knoten und alle Bilder werden sichtbar. Da die Zoomfunktion nur die Entfernungen der Knotenpunkte vergrößert, war die Wahl des Grenzwertes einfach. Es wurde das Zoomlevel gewählt, bei welchem die Entfernungen zwischen den Knotenpunkten ausreichend groß sind, um die Bilder dazwischen anzuordnen.

Zur Anordnung um den Knotenpunkt existieren zwei Methoden. Die Anordnung im Kreis (Abbildung 3.7 auf Seite 13) und die Anordnung im Raster (Abbildung 3.6 auf Seite 13). Bei der Anordnung im Raster werden alle Bilder eines Knotens neben und übereinander angeordnet und das entstandene Raster auf dem Knoten zentriert. Das

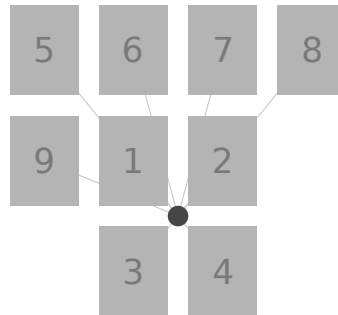


Abbildung 3.6: Anordnung der Bilder im Raster. Der Knotenpunkt liegt im Zentrum des Rasters, welches zeilenweise aufgefüllt wird. Die angegebenen Zahlen geben die Reihenfolge an, in welcher die Bilder hinzugefügt wurden.

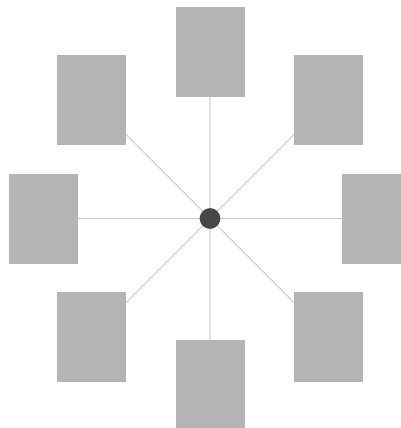


Abbildung 3.7: Anordnung der Bilder in einem Kreis um den Knotenpunkt. Verbindungslinien zeigen von den Bildern zum Knotenpunkt.

Raster wächst dabei von innen nach außen und wird zeilenweise mit Bildern aufgefüllt. Bei der Kreisanordnung wird ein Kreis in ebenso viele Abschnitte unterteilt, wie der Knotenpunkt Bilder besitzt. Jedes Bild wird dann in einem Abschnitt auf dem Kreis platziert. Der Radius steigt mit wachsendem Zoomlevel, so dass sich die Bilder immer weiter auseinander bewegen. Um trotzdem den zugehörigen Knotenpunkt zu einem Bild zu finden, wird eine Verbindungslinie angezeigt.

## 3.5 Maus- und Tastaturinteraktion

Neben der Darstellung der Knoten und Bilder muss die Anwendung auch einfach bedienbar sein. Dabei ist wichtig, dass die Bedienung einfach zu verstehen ist und dem Benutzer nicht ungewohnt erscheint. Die Bewegung der Darstellung erfolgt durch Bewegen der Maus bei gedrückter rechter Maustaste. Mit der linken Maustaste werden Bilder ausgewählt. Ein einfacher Klick auf ein Bild selektiert dieses. Wird die Maus bei gedrückter linker Maustaste bewegt, erzeugt dies eine Auswahlform. Sobald die linke Maustaste wieder losgelassen wird, werden alle Bilder innerhalb der Form selektiert.

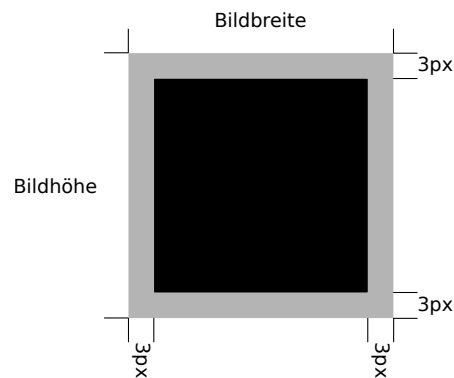


Abbildung 3.8: Zur Vereinfachung der Bildauswahl mit den Auswahlformen, muss von jedem Bild nur der schwarz dargestellte Bereich innerhalb der Auswahl liegen. Die eigentliche Bildgröße ist grau dargestellt.

Es stehen zwei verschiedene Arten von Auswahlformen zur Verfügung. Die Pfadauswahl und die Rechteckauswahl. Wird die Pfadauswahl verwendet, so kann eine Linie um die gewünschten Bilder gezogen werden. Wenn die linke Maustaste losgelassen wird, verbindet sich der Endpunkt mit dem Startpunkt und erzeugt damit die Auswahlform. Die Rechteckauswahl spannt eine rechteckige Form auf. Die Diagonale zwischen dem Startpunkt und der aktuellen Mausposition bestimmt die Größe des Rechtecks. Sobald die Maustaste losgelassen wird, werden die Bilder innerhalb des aktuellen Rechtecks ausgewählt. Um die Auswahl zu vereinfachen, muss von jedem Bild nur ein bestimmter Bereich innerhalb der Auswahlform liegen. Dieser reicht von der Mitte bis 3 Pixel vom Bildrand. (Siehe Abbildung 3.8 auf Seite 14)

Wie viele Bilder ausgewählt werden, hängt von der aktuellen Zoomstufe ab. Ist diese unter dem Grenzwert, werden nur Repräsentanten angezeigt. Wird ein solcher ausgewählt, so werden alle Bilder des Knotens ausgewählt. Ist das Zoomlevel über dem Grenzwert werden alle Bilder angezeigt. Daher wird nur das angeklickte Bild oder die Bilder innerhalb der Auswahlform selektiert.

Mit Hilfe des Reglers in der Kontrollleiste lässt sich die aktuelle Auswahl vergrößern und verkleinern. Dabei wird das Zentrum der aktuellen Auswahl als Mittelpunkt eines Kreises verwendet, dessen Radius man über den Regler beeinflussen kann. Wird der Regler nach rechts bewegt, vergrößert sich der Radius und damit die Auswahl in alle Richtungen. Die Bewegung nach links bewirkt das Gegenteil, bis nur noch das Bild im Mittelpunkt ausgewählt ist.

Das Musrad dient zum Zoomen in der Darstellung. Wenn das Musrad nach oben gedreht wird, vergrößern sich die Abstände zwischen den Knotenpunkten. Wird es nach unten gedreht, verringern sich die Abstände. Die Bilder bleiben dabei immer in der selben Größe. Das ist zwar ungewöhnlich für eine zoombare Oberfläche (ZUI) jedoch durchaus so gewollt. Das Hineinzoomen dient in der Darstellung dazu alle Bilder ab einem Grenzwert sichtbar zu machen. Dabei soll aber das Auswählen der Bilder nicht

schwieriger werden. Wenn die Bilder vergrößert würden, wäre eine größere Auswahlform und damit mehr Aufwand nötig um die Bilder auszuwählen.

Mit der Shift und Alt-Taste können die Funktionen der Maustasten verändert werden. Während die Shift-Taste gedrückt ist, wird eine neue Auswahl der aktuellen hinzugefügt oder von ihr abgezogen. Ist ein Bild der neuen Auswahl bereits ausgewählt, so wird es aus der Auswahl entfernt. Ist es noch nicht selektiert, so wird es zur Auswahl hinzugefügt. Die Funktion verfährt also invertierend. Durch Drücken der Alt-Taste wandelt sich die Auswahlfunktion der linken Maustaste in eine Verschiebefunktion. Während sie gedrückt ist, können die ausgewählten Bilder verschoben werden. Dabei wird unterschieden, ob gerade nur der Repräsentant des Knotens sichtbar ist, oder alle Knoten. Im ersten Fall wird der gesamte Knotenpunkt innerhalb der Darstellung verschoben. Im zweiten Fall verschieben sich nur die gewählten Bilder in Bezug zu ihrem Knotenpunkt.

## 3.6 Visuelle Hilfen

Zur Vereinfachung des Kategorisierens wurden die in Abschnitt 3.3 auf Seite 9 beschriebenen SOM-Darstellungen entwickelt. Zusätzlich gibt es einige unterstützende Funktionen in der Anwendung, welche unabhängig der gewählten SOM-Darstellung sind.

Zur Erleichterung der Navigation werden am Rand der Darstellung rote Kompasslinien angezeigt. Diese Linien zeigen auf noch nicht gelabelte Bilder außerhalb des sichtbaren Bereiches. Sie geben außerdem Auskunft über die Entfernung der Bilder zum Mittelpunkt der Ansicht. Je länger die Linie ist, desto näher ist das Bild am Mittelpunkt der Ansicht. Für bereits gelabelte Bilder wird keine Kompasslinie angezeigt, da die Linien auch den Labelfortschritt darstellen sollen. Inspiriert sind die Kompasslinien aus Computerspielen. Bei diesen wird der Spieler oft mit Hilfe ähnlicher Elemente auf zu erledigende Aufgaben oder zu besuchende Orte hingewiesen.

Durch die Bildstatusfunktion kann jedes Bild entweder den Status „Aktiv“ oder „Inaktiv“ haben. Ungelabelte Bilder haben normalerweise den Status „Aktiv“. Das bedeutet sie sind opak bzw. nicht transparent und können ausgewählt werden (Siehe Abbildung 3.9 auf Seite 16). Sobald einem Bild ein Label zugewiesen wird, ändert sich dessen Status automatisch zu „Inaktiv“. Es wird halbtransparent dargestellt und kann nicht mehr ausgewählt werden (Siehe Abbildung 3.10 auf Seite 16). Über die Labelleiste kann der Status einer Kategorie bzw. eines Labels geändert werden werden. Dafür ist vor jedem Labelknopf ein Kontrollkästchen, welches den aktuellen Status anzeigt und beim Anklicken wechselt. Statuswechsel beziehen sich immer auf alle Bilder mit dem selben Label. Um den Status von allen ungelabelten Bildern zu wechseln, steht ein extra Kontrollkästchen zur Verfügung.



Abbildung 3.9: Beispiel für ein aktives Bild. Es ist nicht transparent und kann ausgewählt werden.



Abbildung 3.10: Beispiel für ein inaktives Bild. Es ist transparent und kann nicht ausgewählt werden. Weiterhin wurde es beim Labeln mit der Labelfarbe eingefärbt.

Die Deaktivierung der Bilder soll einerseits die Übersichtlichkeit erhöhen. Denn die halb transparenten Bilder fallen weniger auf und der Benutzer kann sich besser auf die übrigen Bilder konzentrieren. Zum Anderen soll das Auswählen von Bildern vereinfacht werden. Denn die Auswahlformen können deaktivierte Bilder überdecken, ohne diese mit auszuwählen. Besonders bei eng aneinander liegenden Bildern verschiedener Kategorien vereinfacht dies das Arbeiten.

Falls dem Benutzer die Deaktivierung noch nicht ausreicht, können alle gelabelten Bilder auch komplett ausgeblendet werden. Dazu befindet sich ein extra Kontrollkästchen in der Labelleiste.



## 4 Evaluation

### 4.1 Testaufbau und -ablauf

Mit einer Benutzerstudie sollte festgestellt werden, ob die Vorsortierung mit Hilfe einer selbstorganisierenden Karte wirklich hilfreich ist. Die folgenden Punkte sollten im Test überprüft werden:

- Vergleich zwischen Vorsortierung mit SOM und ohne Vorsortierung
- Vergleich der normalen SOM-Darstellung mit der verschobenen SOM-Darstellung und der farbigen SOM-Darstellung
- Bewertung der Bildstatusfunktion
- Bewertung der Kompasslinien
- Vergleich zwischen Auswahl mit Pfad und mit Rechteck

Zusätzlich sollten Wünsche für neue Funktionen gesammelt werden.

Zum Testen wurde eine modifizierte Version der Anwendung benutzt. Diese zeigte weder eine Menüleiste noch die Funktion zum Erstellen und Ändern von Labels. Zur einfacheren Auswertung sollten die Tester keine eigenen Kategorien vergeben, sondern vorgegebene Labels verwenden. Weiterhin konnten mit der Evaluationsversion verschiedene Nutzungsdaten aufgezeichnet werden. Diese Daten sollten, zusammen mit den Aussagen der Tester, die gerade genannten Punkte klären.

Folgende Daten wurden zu jedem Test aufgezeichnet:

- Gesamtzeit des Labelvorgangs
- Wie lange wurde die Darstellung verschoben?
- Wie oft wurde die Darstellung verschoben?
- Verwendungszeit des Pfadauswahlwerkzeuges
- Verwendungszeit des Rechteckauswahlwerkzeuges
- Anzahl der Selektionen durch Mausklick
- Wie oft wurde gezoomt?
- Wie oft wurde zwischen Status „Aktiv“ und „Inaktiv“ gewechselt?

Außerdem wurde gespeichert welche Label welchen Bildern zugeordnet wurden. Damit sollten Fehleranfälligkeiten der Darstellungen gemessen werden.

Für den Test standen vier Bilddatensätze zur Verfügung. Die ersten beiden bestanden aus 5 und 19 Objekten der „Columbia Object Image Library (COIL-100)“ [Nene et al., 1996] [1] (Siehe Abbildung 4.1 auf Seite 18). Diese Objekte werden häufig für



Abbildung 4.1: Drei Objekte aus dem COIL-100-Datensatz. Die Bilder wurden in zwei Testdatensätzen verwendet. Ein kleinerer mit 5 Objekten (432 Bilder) und ein größerer mit 19 Objekten (857 Bilder). Von jedem Objekt wurde eine andere Anzahl Bilder, aus verschiedenen Winkeln, verwendet.



Abbildung 4.2: Drei Bilder aus dem Fensterdatensatz, welcher 666 Bilder enthielt. Die Bilder stammen von drei verschiedenen Fenstern an der selben Wand. Unterschieden werden sollte, ob das Fenster geöffnet oder geschlossen ist.

Objekterkennungstests verwendet und sind leicht zu unterscheiden. Von jedem Objekt gibt es Photos aus verschiedenen Winkeln. Im Test wurde für jedes der 19 Objekte des größeren Datensatzes eine andere Anzahl Bilder verwendet um einige Objekte schwieriger auffindbar zu machen. Insgesamt befanden sich im Datensatz mit 5 Objekten 432 Bilder und im Datensatz mit 19 Objekten 857 Bilder. In der weiteren Auswertung wird der größere Datensatz als COIL-100-Datensatz bezeichnet.

Ein weiterer Datensatz bestand aus Webcam-Photos dreier Fenster im Institut. Diese waren teils offen, teils geschlossen und in verschiedenen Beleuchtungsbedingungen aufgenommen. Die Bilder waren so beschnitten, dass nur das Fenster sichtbar ist. Im Datensatz waren insgesamt 666 Bilder enthalten. (Siehe Abbildung 4.2 auf Seite 18).

Der dritte Datensatz bestand aus Webcam-Aufnahmen eines Parkplatzes der Universität Paderborn, bestehend aus 153 Bildern [2]. Wieder wurden Aufnahmen verschiedener Tageszeiten verwendet und ein Bildausschnitt gewählt, welcher nur den Bereich zeigt, der für das Kategorisieren wichtig ist. (Siehe Abbildung 4.3 auf Seite 19)



Abbildung 4.3: Drei Bilder aus dem Parkplatzdatensatz. Die Bilder stammen von einer Webcam an der Universität Paderborn [2]. Unterschieden werden sollte, ob der Parkplatz leer, mittel oder voll ist.

Datensatz	SOM-Darstellung	SOM-Darstellung mit farbigen Hexagons	SOM-Darstellung mit verschobenen Knotenpunkten	Unsortierte Darstellung
COIL-100	x	x		x
Fenster	x	x		x
Parkplatz	x		x	x

Tabelle 4.1: Für jeden Datensatz wurden 3 Darstellungen getestet. die getesteten Darstellungen sind mit einem x markiert. Zu Beginn der Tests sollten alle 4 Darstellungen für jeden Datensatz getestet werden. Dies musste jedoch geändert werden, da die Tests sonst zu lange gedauert hätten.

Die Aufgabenstellung war für alle Datensätze die selbe: jedes Bild soll einem der vorgegebenen Label zugewiesen werden. Dafür gab es beim COIL-100-Datensatz ein Label für jedes Objekt, beim Fensterdatensatz ein Label für „offen“ und eins für „geschlossen“ und beim Parkplatzdatensatz die Label „leer“, „mittel“ und „voll“. Zu Beginn der Evaluation war geplant, dass jeder Datensatz vier mal gelabelt wird. Mit der Standard-SOM-Darstellung, der unsortierten Darstellung, der SOM-Darstellung mit verschobenen Knotenpunkten und der SOM-Darstellung mit farbigen Hexagons. Dies erwies sich aber als sehr zeitaufwändig, weshalb die Anzahl auf 3 Darstellungen pro Datensatz reduziert wurde. Da die SOM-Darstellung mit verschobenen Knotenpunkten als unübersichtlichste SOM-Darstellung erschien, wurde diese beim COIL und Fensterdatensatz nicht getestet. Beim Parkplatzdatensatz wurde dafür die SOM-Darstellung mit farbigen Hexagons nicht getestet. (Siehe Tabelle 4.1 auf Seite 19)

Der COIL Datensatz mit 5 Objekten wurde genutzt um zu testen, ob die Bildstatusfunktion einen Vorteil bringt. Jeder Tester musste den Datensatz zweimal labeln. Einmal konnten die Status „Aktiv“ und „Inaktiv“ auf die Bilder angewendet werden. Diese wurden dann entweder transparent und nicht selektierbar („Inaktiv“) oder opak und selektierbar („Aktiv“). Im anderen Durchlauf war diese Möglichkeit deaktiviert und alle Bilder blieben ständig im Status „Aktiv“. Die zu vergebenden Label waren wieder vorgegeben und bestanden aus den Bezeichnungen der Objekte.

Es ergab sich also folgender Testablauf:

1. Tutorial mit einem COIL Datensatz, bestehend aus 5 Objekten. Keine Zeitmessung, diente nur zur Erklärung der Programmfunktionen und der verschiedenen Darstellungen.
2. COIL Datensatz mit 5 Objekten (nicht identisch mit dem Tutorialdatensatz). Getestete Darstellungen: Bildstatusfunktion an, Bildstatusfunktion aus
3. COIL Datensatz mit 19 Objekten. Getestete Darstellungen: unsortiert, SOM, SOM mit farbigen Hexagons
4. Fensterdatensatz. Getestete Darstellungen: unsortiert, SOM, SOM mit farbigen Hexagons
5. Parkplatzdatensatz. Getestete Darstellungen: unsortiert, SOM, SOM

Um Lerneffekte bei den Testern nachträglich herausrechnen zu können, wurden 12 Testgruppen erstellt, jede mit einer anderen Reihenfolge der Darstellungen.

Zu jeder Darstellung wurden den Testern Fragen gestellt. Es sollte zuerst die Darstellung mit eigenen Worten bewertet werden. Danach sollte der Eindruck zur Schwierigkeit und Geschwindigkeit auf einer Skala von sehr leicht bis sehr schwer, bzw. sehr schnell bis sehr langsam angegeben werden. Weiterhin wurden die Tester gebeten während des Tests ihre Eindrücke zu formulieren, welche auch notiert wurden. Im Anschluss an die Tests eines Datensatzes sollte außerdem die beste der getesteten Darstellungen gewählt werden.

Als Abschluss der Tests folgte eine weitere Befragung, bei welcher folgende Fragen zur Anwendung beantwortet werden sollten:

- Welche der gezeigten Darstellungen gefällt Ihnen am besten? Warum?
- Hilft die Visualisierung der U-Matrix bei der Zuordnung gleicher Bilder zu einer Kategorie?
- Wie empfinden Sie das Verhalten der Oberfläche beim hineinzoomen?
- Wie finden Sie die Darstellung mehrerer Bilder durch ein Bild?
- Welche Eingabeoption fanden Sie am besten? Wie bewerten Sie die einzelnen Eingabeoptionen (Lasso, Mausklick, Rechteck)?
- Hat der Kompass bei der Orientierung in den großen Datensätzen geholfen?
- Empfinden Sie eine der vorgestellten Darstellungen als motivierend?
- Können Sie sich vorstellen große Datensätze damit zu labeln? Warum? Warum nicht?
- Wie finden Sie die Oberflächen allgemein?
- Hat Sie etwas gestört?

Reihenfolge	SOM-Darstellung	SOM-Darstellung mit farbigen Hexagons	SOM-Darstellung mit verschobenen Knotenpunkten	Unsortierte Darstellung
1. Versuch	10	6	4	16
2. Versuch	16	12	4	4
3. Versuch	10	6	4	16

Tabelle 4.2: Anzahl der vorhandenen Testwerte pro Darstellung und Versuch. Die Einteilung der Werte nach Versuchen soll die vorhandenen Lerneffekte ausgleichen. So können alle Darstellungen eines Versuchs untereinander verglichen werden. Um die normale SOM-Darstellung und die Hexagon-SOM-Darstellung direkt vergleichen zu können, wurden beide immer direkt nacheinander, in verschiedenen Reihenfolgen, getestet. Daher kommt es zu den unterschiedlichen Werteanzahlen pro Versuch. Im Vergleich der SOM-Darstellungen mit der unsortierten Darstellung werden die Werte aus dem 2. Versuch nicht verwendet, da nur 4 Werte für die unsortierte Darstellung existieren.

## 4.2 Testauswertung

### 4.2.1 Vergleich zwischen sortierter und unsortierter Darstellung

Um herauszufinden, ob die Vorsortierung einen Vorteil beim Kategorisieren bringt, sollen alle sortierten Darstellungen mit der unsortierten Darstellung verglichen werden. Dazu wird der Mittelwert aus allen Zeiten aller SOM-Darstellungen und aus allen Zeiten der unsortierten Darstellung gebildet.

Da leider kein Weg gefunden wurde, die Lerneffekte aus den Ergebnissen herauszurechnen, wurden die Ergebnisse in Gruppen unterteilt. Jede Darstellung konnte entweder als erster, zweiter oder dritter Versuch an der Reihe sein.

Da die Anordnung der Bilder in der normalen SOM-Darstellung identisch zur SOM-Darstellung mit farbigen Hexagons ist, sollten beide direkt miteinander verglichen werden. Dafür wurden beide Darstellungen immer direkt nacheinander, in verschiedenen Reihenfolgen, getestet. Die sich daraus ergebende unterschiedliche Anzahl von Tests pro Darstellung ist in Tabelle 4.2 auf Seite 21 vermerkt. Da die Werteanzahl der unsortierten Darstellung im zweiten Versuch sehr gering ist, werden in diesem Vergleich nur Ergebnisse verwendet, bei welchen die Darstellungen als erster oder dritter Versuch an der Reihe waren. Wie in Tabelle 4.2 zu sehen gibt es in Versuch 1 und 3 jeweils 20 Werte für die SOM-Darstellungen und 16 Werte für die unsortierte Darstellung. Aus diesen werden die Mittelwerte zum Vergleich gebildet. An den Werten selbst wurden keine Veränderungen vorgenommen. Falls Ausreißer das Ergebnis beeinflussen, ist dies an den jeweiligen Diagrammen vermerkt. In den Diagrammen weisen Pfeile symbolisch auf Ausreißer hin. Sie weisen in die Richtung, in welche der Wert verändert werden müsste.

Zuerst soll die Zeit verglichen werden, welche die Tester durchschnittlich zum Katego-



Abbildung 4.4: Dargestellt ist die Zeit, welche eine Testperson durchschnittlich für das Labeln eines Datensatzes benötigte. Es wird unterteilt, ob die Darstellung als erste oder dritte Darstellung eines Datensatzes bearbeitet wurde. Die unsortierte Darstellung wurde zu selten als zweiter Versuch verwendet, darum sind für diesen Fall keine Werte angegeben. Die dargestellten Werte sind Mittelwerte aus 20 Werten mit Vorsortierung und 16 Werten ohne Vorsortierung. Die Werte mit Vorsortierung sind zusammengefügt aus allen Zeiten der normalen SOM-Darstellung (10 Zeiten), der verschobenen SOM-Darstellung (4 Zeiten) und der SOM-Darstellung mit farbigen Hexagons (6 Zeiten).

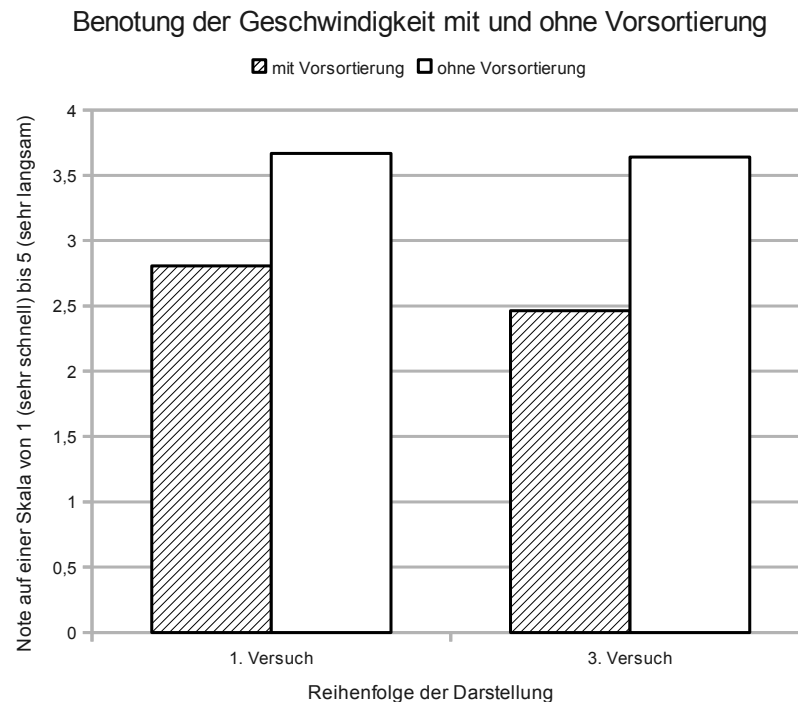


Abbildung 4.5: Dargestellt ist die Selbsteinschätzung der Testpersonen zur Labelgeschwindigkeit. Nach dem Labeln wurden die Tester um eine Einschätzung ihrer Geschwindigkeit auf einer Skala von „sehr schnell“ bis „sehr langsam“ gebeten. Diese Angaben werden hier in numerischen Werten angezeigt. „Sehr schnell“ entspricht einem Wert von 1, „sehr langsam“ einem Wert von 5. Die Werte mit Vorsortierung sind wieder der Mittelwert aus allen Zeiten der normalen SOM-Darstellung, der verschobenen SOM-Darstellung und der SOM-Darstellung mit farbigen Hexagons. Die Anzahl der verwendeten Werte ist in Tabelle 4.2 auf Seite 21 angegeben.

risieren der Datensätze benötigten. In Abbildung 4.4 auf Seite 22 ist zu sehen, dass die vorsortierten Darstellungen ein bis drei Minuten schneller bearbeitet wurden, als die unsortierte Darstellung. Auch der Lerneffekt war bei der vorsortierten Darstellung größer. Denn aus dem 1. und 2. Versuch waren die Bilder des Datensatzes und die zuzuordnenden Label bereits bekannt. Trotzdem ist die Geschwindigkeitssteigerung bei der unsortierten Darstellung gering. Darum kann angenommen werden, dass die unsortierte Darstellung schwieriger zu labeln ist.

Der höhere Schwierigkeitsgrad der unsortierten Darstellung lässt sich auch in Diagramm 4.5 auf Seite 23 ablesen. In der Befragung, welche nach jeder Darstellung eines Datensatzes durchgeführt wurde, sollten die Tester eine Einschätzung ihrer Geschwindigkeit treffen. Diese wurde auf einer Skala mit den Werten: „sehr schnell“, „eher schnell“, „mittel“, „eher langsam“ und „sehr langsam“ eingetragen. Im Diagramm sind Werte in der selben Art wie Schulnoten mit Zahlen von 1 („sehr schnell“) bis 5 („sehr langsam“) eingetragen.

Im Diagramm 4.6 auf Seite 24 sind die Empfindungen im Bezug auf den Schwierig-



Abbildung 4.6: Dargestellt ist die Einschätzung der Testpersonen zur Schwierigkeit bzw. Aufwendigkeit des Labelvorgangs. Diese sollte auf einer Skala von „sehr einfach“ bis „sehr schwer“ angegeben werden. Im Diagramm entspricht die Skala numerischen Werten von 1 („sehr einfach“) bis 5 („sehr schwer“). Die Werte mit Vorsortierung sind wieder der Mittelwert aus allen Zeiten der normalen SOM-Darstellung, der verschobenen SOM-Darstellung und der SOM-Darstellung mit farbigen Hexagons. Die Anzahl der verwendeten Werte ist in Tabelle 4.2 auf Seite 21 angegeben.



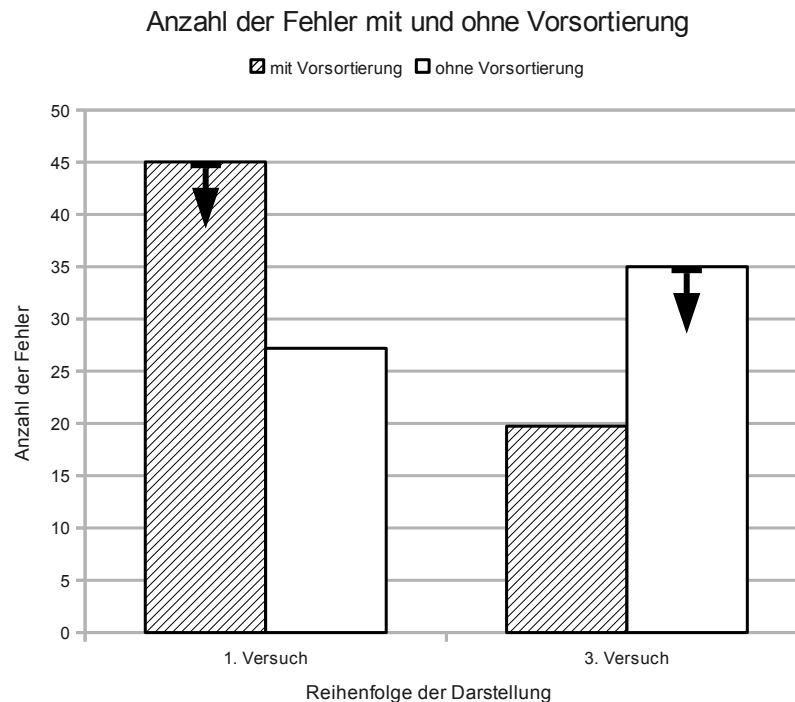


Abbildung 4.7: Dargestellt ist die durchschnittliche Anzahl von fehlerhaft gelabelten Bildern im 1. und 3. Versuch. Die Korrektheit der vergebenen Label wurde mit Hilfe einer Vergleichsliste festgestellt. In dieser Liste wurden die richtigen Labels zu den jeweiligen Bildern vermerkt. Im Laufe der Tests zeigte sich jedoch, dass die Datensätze Fenster und Parkplatz mit so einer Liste weniger gut auszuwerten sind. Denn die Einschätzungen zum Zustand eines Bildes waren sehr subjektiv. Auch beim COIL-100-Datensatz kam es zu fehlerhaften Labelzuordnungen, auf Grund von Ungenauigkeiten bei den vorgegebenen Labelnamen. Die beiden Probleme führten zu einigen Ausreißern. Im ersten Versuch müsste die Fehleranzahl der vorsortierten Darstellung geringer sein, im zweiten Versuch die der unsortierten. Die Ausreißer sind durch Pfeile symbolisiert, welche in die Richtung weisen, in welche der Wert verändert werden müsste.

keitsgrad direkt abzulesen. Die Testpersonen konnten wieder auf einer fünfteiligen Skala ihre Meinung angeben, welche im Diagramm durch die Werte 1 („sehr einfach“) bis 5 („sehr schwer“) sichtbar ist. Ähnlich wie bei der Benotung der Geschwindigkeit und der gemessenen Geschwindigkeit fallen die Ergebnisse positiv für die Vorsortierung aus. Interessant zu bemerken ist, dass die unsortierte Darstellung beim 3. Versuch des Datensatzes schlechter bewertet wurde, als beim 1. Versuch. Die Testpersonen schienen die vorsortierten Darstellungen, welche sie vorher gesehen haben zu bevorzugen.

Ein weiterer Vergleichspunkt zwischen den beiden Darstellungsarten ist die Korrektheit der vergebenen Labels. Auch wenn die vorsortierte Darstellung schneller zu labeln ist, kann sie nicht überzeugen wenn mit ihr zu viele Fehler gemacht werden. Dabei sind keine Programmfehler gemeint, welche die Bilder unabhängig vom Nutzer falsch labeln, sondern vom Benutzer versehentlich falsch gelabelte Bilder. Diese können zum Beispiel auftreten, wenn die Darstellung nicht übersichtlich genug ist und einzelne Bilder aus

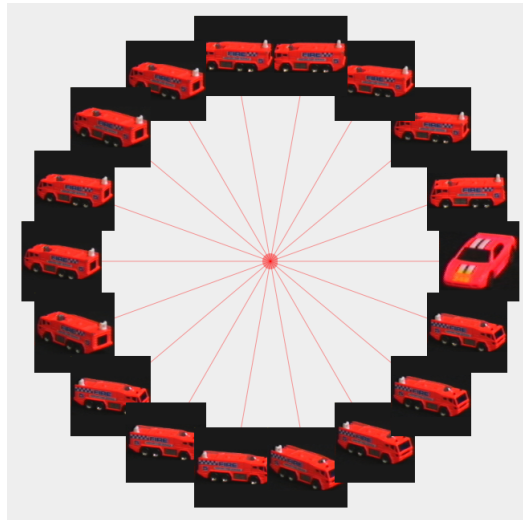


Abbildung 4.8: Dieses, aus dem COIL-100 Datensatz stammende Beispiel zeigt ein schwierig zu findendes, fehlerhaft zugeordnetes Bild eines Knotens. Das einzelne Auto (mitte, rechts) müsste das Label „Auto rot“ erhalten, die übrigen „Feuerwehrauto“. Da sich das Bild nur wenig von den anderen unterscheidet, ist es sehr leicht zu übersehen und wurde mehrmals falsch gelabelt.

Versehen zusammen mit einer anderen Bildgruppe ausgewählt werden (Siehe Abbildung 4.8 auf Seite 26).

Um die Korrektheit zu messen, wurden die zugewiesenen Label gespeichert und eine Liste mit korrekten Labels zum Vergleich erstellt. Theoretisch sollte es durch einen einfachen Abgleich mit der korrekten Liste möglich sein die Anzahl der Fehler zu finden. Es stellte sich jedoch heraus, dass der Fenster- und der Parkplatzdatensatz nur bedingt dafür geeignet sind. Das Problem ist, dass die Beurteilung eines Objektzustandes vom jeweiligen Betrachter abhängig ist. Einige Tester schätzten den Parkplatz als voll ein, obwohl andere Testpersonen das Label „mittel“ verwendeten. Ein weiteres Problem bestand beim COIL-100-Datensatz. Viele Tester hatten mit den vorgegebenen Kategorien Probleme, da sie die zugehörigen Objekte nicht mit den selben Worten beschreiben würden. So wurde eine türkisfarbene Limonadendose oft als „grüne Dose“ kategorisiert, obwohl das Label für eine andere Dose vorgesehen war. Diese Probleme führten zu einigen Ausreißern unter den aufgezeichneten Werten und verfälschen damit Diagramm 4.7 auf Seite 25. Eigentlich müssten die beiden Darstellungen näher beieinander liegen. Die vorsortierte Darstellung im ersten Versuch und die unsortierte Darstellung im 3. Versuch müssten beide einen geringeren Wert haben. Durch Beobachtungen während der Arbeit der Testpersonen lässt sich aber dennoch eine Aussage über die Korrektheit der beiden Darstellungen treffen. Die unsortierte Darstellung hatte, vor allem bei kleinen Datensätzen, weniger Fehler. Durch die unübersichtliche Anordnung, wurden die Testpersonen scheinbar gezwungen sehr konzentriert zu arbeiten. Meistens haben die Tester Bilder übersehen, welche einzeln zwischen großen Gruppen anderer Bilder lagen. Bei den vorsortierten Darstellungen kam es öfter zu falsch kategorisierten Bildern.

Den Testpersonen wurde scheinbar der Eindruck vermittelt, die Bilder eines Knotens entsprächen immer den Repräsentanten. Darum wurden die restlichen Bilder eines Knotens oft nur kurz überflogen. Bilder zu finden, welche bei der SOM Berechnung dem falschen Knoten zugeordnet wurden, erschien schwierig.

Zusammenfassend können die vorsortierten Darstellungen unter Vorbehalt als vorteilhaftere Darstellung bezeichnet werden. Sowohl gemessen, als auch in der persönlichen Wahrnehmung der Tester waren sie schneller als die unsortierte Darstellung. Auch wurden sie als übersichtlicher empfunden. Einzig bei der Korrektheit schienen die vorsortierten Darstellungen zurückzubleiben. Dies kann durchaus an der Verwendung der Repräsentanten liegen. Diese müssten durch etwas anderes ersetzt werden, was die Summe der Bilder eines Knotens besser abbildet. Es muss vor allem sichtbar werden, ob die Bilder des Knotens wirklich alle gleich sind.

## 4.2.2 Vergleich der verschiedenen SOM-Darstellungen

In diesem Abschnitt soll festgestellt werden, welche Darstellung der self organizing map die beste ist. Kriterien dafür sind zum einen die zum Labeln benötigte Zeit und die Anzahl der Fehler. Zum anderen die Meinung der Testpersonen. Diese wurden nach jedem Datensatz befragt, welche Darstellung die beste für diesen Datensatz war. Die Stimmen wurden gezählt und die Darstellung mit den meisten Stimmen ist die beliebteste. Die verschiedenen Darstellungen waren: die normale SOM-Darstellung, die SOM-Darstellung mit verschobenen Knotenpunkten und die SOM-Darstellung mit farbigen Hexagons als Hintergrund. Der Vergleich der drei SOM-Darstellungen erfolgt fast identisch zum Vergleich zwischen den vorsortierten Darstellungen und der unsortierten Darstellung. Es werden immer nur zwei Darstellungen miteinander verglichen. Die Werte sind wieder nach Versuchen unterteilt, da die Einflüsse des Lerneffekts nicht entfernt werden konnten.

Zu erst soll die normale SOM-Darstellung mit der verschobenen SOM-Darstellung verglichen werden. Da die SOM-Darstellung mit verschobenen Knoten nur im Parkplatzdatensatz getestet wurde, werden für den Vergleich nur Werte aus diesem Datensatz verwendet. In Diagramm 4.9 auf Seite 28 werden die Labelzeiten der normalen SOM-Darstellung und der SOM-Darstellung mit verschobenen Knoten verglichen. Die Werte jedes Versuchs sind Mittelwerte der Zeiten von 4 Testpersonen. Zwei Ausreißer verfälschen die Ergebnisse. Der erste erhöht die Zeit der verschobenen SOM-Darstellung in Versuch 1. Jedoch müsste die Zeit auch ohne den Ausreißer über der der SOM-Darstellung liegen. Der zweite Ausreißer erhöht die Zeit der normalen SOM-Darstellung in Versuch 3. Diese müsste unter der Zeit der verschobenen SOM-Darstellung liegen. Es ist zu sehen, dass die Zeitunterschiede eher gering sind. Das liegt vor allem daran, dass der Parkplatzdatensatz nur 153 Bilder enthielt und die benötigte Zeit damit sehr kurz war. Bei der normalen SOM-Darstellung sind die Zeiten jedoch ein wenig kürzer.

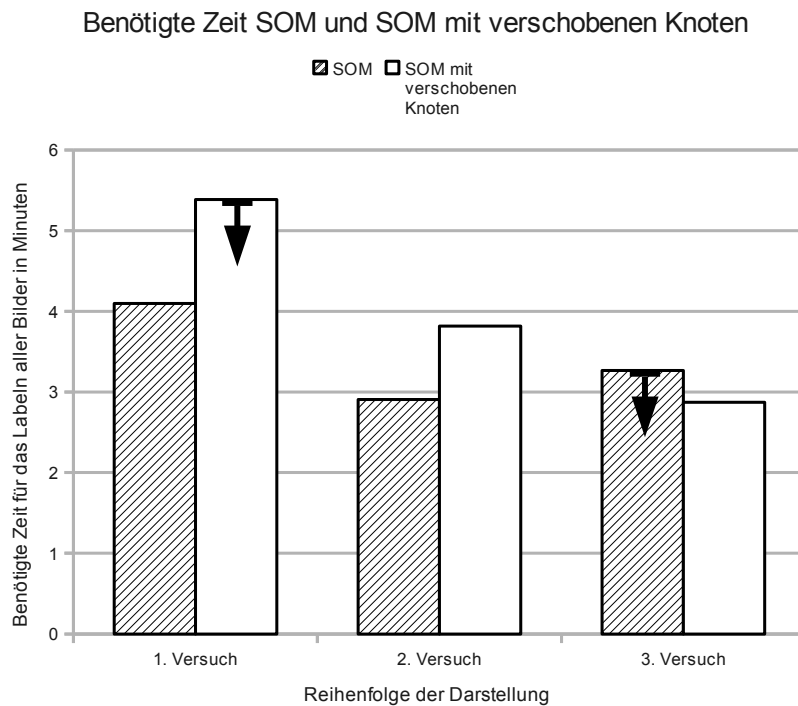


Abbildung 4.9: Vergleich der durchschnittlichen Labelzeiten im Parkplatzdatensatz. Jeder Versuch besteht aus dem Mittelwert von 4 Testern. Ein Ausreißer bei Versuch 1 erhöht den Wert der SOM-Darstellung mit verschobenen Knoten. Die Zeit müsste jedoch trotzdem höher sein als die der normalen SOM-Darstellung. Bei Versuch 3 sorgt ein Ausreißer dafür, dass die normale SOM-Darstellung einen höheren Zeitwert besitzt als die verschobene SOM. Dieser müsste unter der verschobenen SOM liegen. Die Ausreißer sind durch Pfeile symbolisiert, welche in die Richtung weisen, in welche der Wert verändert werden müsste.

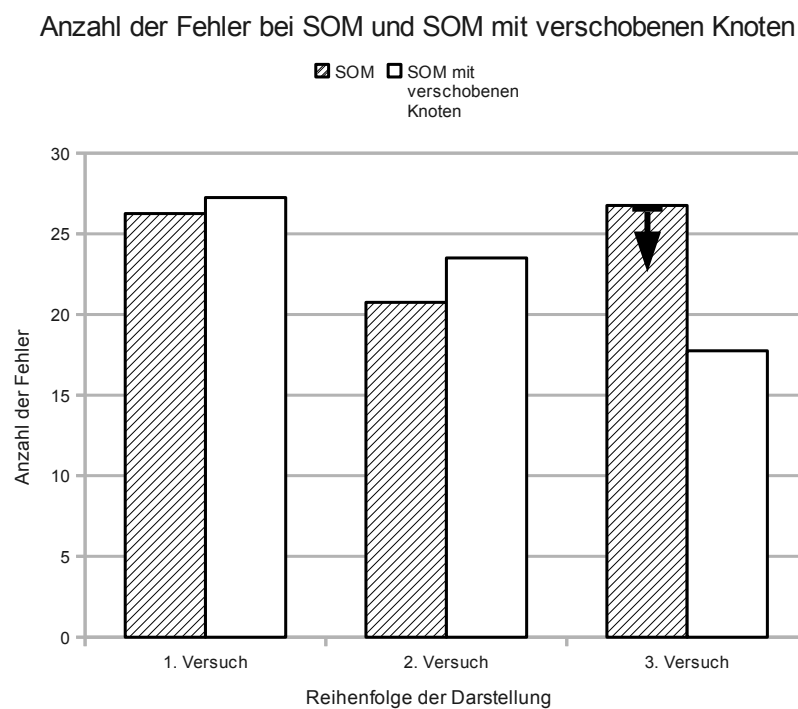


Abbildung 4.10: Durchschnittliche Anzahl der Fehler bei normaler SOM-Darstellung und SOM-Darstellung mit verschobenen Knotenpunkten. Die Werte sind Mittelwerte aus den Fehlerzahlen von 4 Testern. Sie stammen ausschließlich aus dem Parkplatzdatensatz. Ein kleiner Ausreißer erhöht die Anzahl der Fehler bei der normalen SOM-Darstellung in Versuch 3. Jedoch wäre die Anzahl auch ohne den Ausreißer höher als die der verschobenen SOM.

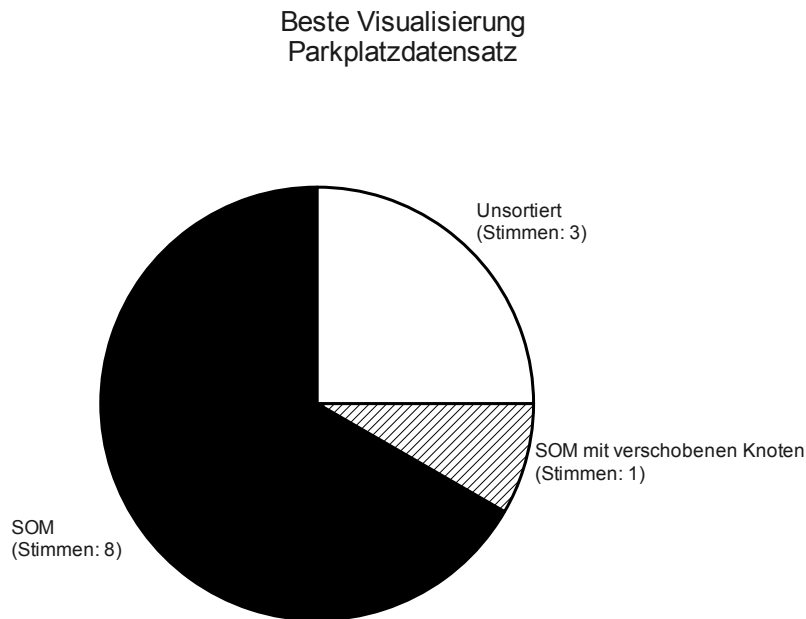


Abbildung 4.11: Bewertung aller Darstellungen des Parkplatzdatensatzes. Die Stimmen wurden im Anschluss an die Labelvorgänge des Datensatzes abgegeben. Jeder Tester konnte eine der drei Darstellungen als beste für diesen Datensatz auswählen.

Die Anzahl der Fehler wird in Diagramm 4.10 auf Seite 29 dargestellt. Die Werte sind wieder das Mittel aus den Ergebnissen von 4 Testpersonen. Ein kleiner Ausreißer erhöht bei Versuch 3 die Anzahl der Fehler der normalen SOM-Darstellung. Jedoch wäre die Anzahl auch ohne den Ausreißer höher als die der SOM-Darstellung mit verschobenen Knoten. Es ist aber zu sehen, dass die Fehlerzahlen sich nicht so sehr unterscheiden. Weiterhin ist anzumerken, dass die Fehleranzahl durch die Art der Labelaufgabe beeinflusst wurde. Die Tester sollten einschätzen ob der Parkplatz voll, mittel oder leer war. Die Entscheidung zwischen leer und mittel und zwischen mittel und voll fiel von Tester zu Tester verschieden aus. Die Fehleranzahl müsste deswegen insgesamt verringert werden.

Im Gegensatz zu den Zeit- und Fehlervergleichen ist beim Meinungsvergleich in Diagramm 4.11 auf Seite 30 ein klarer Unterschied erkennbar. Von den 12 Testern empfanden 8 die normale SOM-Darstellung als am besten geeignet für den Parkplatzdatensatz. Nur 1 Stimme ging an die SOM Darstellung mit verschobenen Knoten. Damit schnitt diese sogar schlechter ab, als die unsortierte Darstellung. Das liegt aber vor allem an der geringen Größe des Datensatzes. Bei den Befragungen wurde die verschobene SOM Darstellung aber oft als unübersichtlich empfunden, weil sich Bilder häufig überlappten.

Die normale SOM-Darstellung scheint im Vergleich zur SOM-Darstellung mit verschobenen Knoten geringfügig schneller und deutlich beliebter zu sein. In Bezug auf Labelfehler

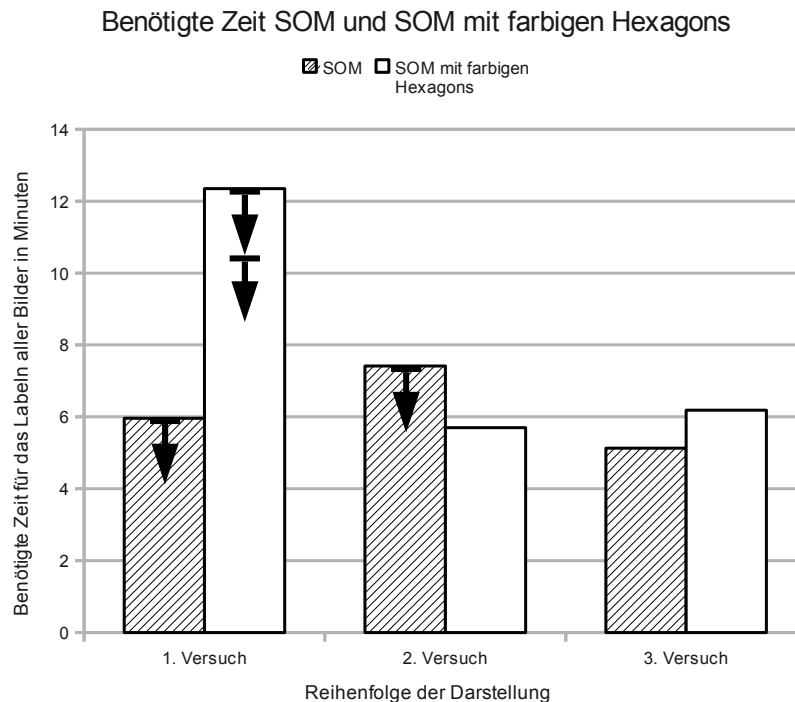


Abbildung 4.12: Dargestellt ist ein Vergleich zwischen den Labelzeiten der normalen SOM-Darstellung und der SOM-Darstellung mit farbigen Hexagons. Die Ergebnisse sind Mittelwerte aus den Zeiten der Datensätze COIL-100 und Fenster. Der erste und der dritte Versuch bildet sich jeweils aus 6 Testergebnissen, der zweite Versuch aus 12. Beim ersten Versuch gibt es zwei Ausreißer bei der Hexagon-SOM-Darstellung und einen bei der normalen SOM-Darstellung. Der Wert der Hexagon-SOM-Darstellung müsste deutlich geringer sein, jedoch trotzdem größer als der der normalen SOM-Darstellung. Im zweiten Versuch wird das Ergebnis der normalen SOM durch einen Ausreißer beeinflusst, aber nur wenig. Der Wert müsste ein bisschen geringer sein.

lässt sich kein entscheidender Unterschied feststellen. Zur Verbesserung der verschobenen SOM-Darstellung wäre es sehr hilfreich an deren Übersichtlichkeit zu arbeiten. Wenn sich weniger oder keine Bilder überlappen, könnte die Darstellung einer weiteren Evaluation unterzogen werden.

Für den Vergleich der normalen SOM-Darstellung und der SOM-Darstellung mit farbigen Hexagons stehen die Ergebnisse der Datensätze COIL-100 und Fenster zur Verfügung. Die Ergebnisse teilen sich wie folgt in die 3 Versuche auf: jeweils 6 Ergebnisse fallen auf den ersten und den dritten Versuch, während 12 Ergebnisse dem zweiten Versuch zugeordnet werden. Das Diagramm 4.12 auf Seite 31 zeigt die mittleren Labelzeiten der normalen SOM-Darstellung und der SOM-Darstellung mit farbigen Hexagons. Im ersten Versuch erhöhen 2 Ausreißer den Wert der Hexagon-SOM-Darstellung und ein weiterer Ausreißer den Wert der normalen SOM-Darstellung. Der Wert der Hexagon-SOM-Darstellung wäre ohne die Ausreißer deutlich geringer, aber trotzdem noch größer als der der normalen SOM. Im zweiten Versuch müsste der Wert der normalen SOM

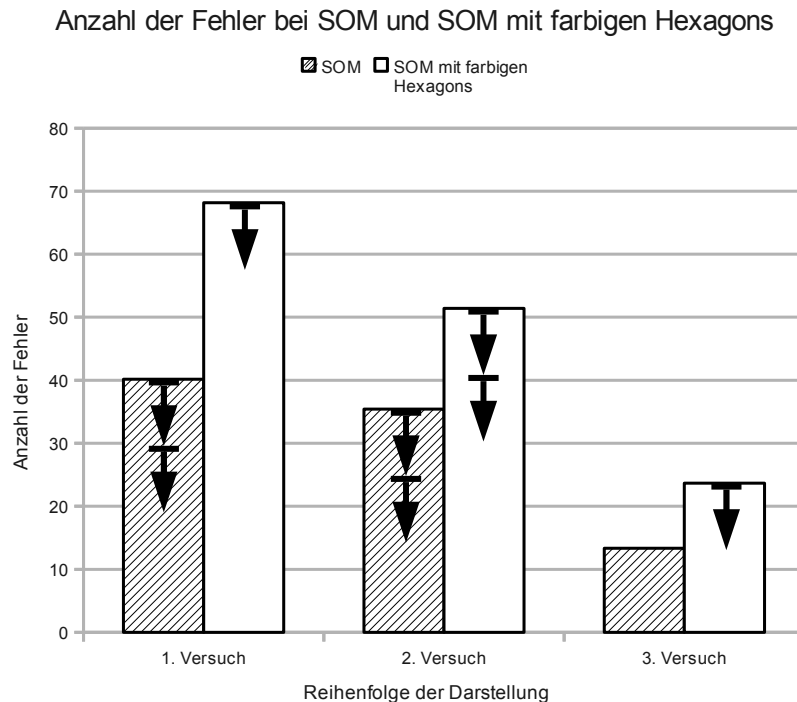


Abbildung 4.13: Durchschnitt der Fehleranzahl bei normaler SOM-Darstellung und SOM-Darstellung mit farbigen Hexagons. Der erste und dritte Versuch bestehen aus jeweils 6 Zeilen, der zweite Versuch aus 12. Alle Zeilen stammen aus dem COIL-100- und dem Fensterdatensatz. Im ersten Versuch wird das Ergebnis der normalen SOM durch zwei Ausreißer nach oben verfälscht, das Ergebnis der Hexagon SOM durch einen Ausreißer. Beide Darstellungen hätten eigentlich deutlich weniger Fehler und die normale SOM-Darstellung läge mit ihrem Wert über dem der Hexagon-SOM-Darstellung. Im zweiten Versuch verfälschen bei beiden Darstellungen zwei Ausreißer das Ergebnis. Ohne diese lägen die Werte deutlich niedriger, die normale SOM-Darstellung läge aber weiterhin vorn. Im dritten Versuch existiert ein kleinerer Ausreißer bei der SOM-Darstellung mit farbigen Hexagons. Eigentlich wäre der Wert dieser Darstellung nahezu gleich auf mit der normalen SOM-Darstellung.

geringer sein. Auch hier erhöht ein Ausreißer das Ergebnis, jedoch nicht sehr stark. Unter Berücksichtigung der Ausreißer, lässt sich sagen, dass die Zeitunterschiede beim Labeln der beiden Darstellungen nicht gravierend sind. Bis auf die Werte beim ersten Versuch liegen die Ergebnisse ähnlich nah beieinander wie im Vergleich der normalen SOM-Darstellung mit der verschobenen SOM-Darstellung.

Der Vergleich der Fehlerzahlen wird durch eine große Anzahl Ausreißer beeinflusst. Im ersten Versuch verfälschen zwei Ausreißer die normale SOM-Darstellung und ein weiterer die SOM-Darstellung mit farbigen Hexagons. Dies bewirkt, dass die Fehlerzahlen stark ansteigen und der Wert der Hexagon-SOM-Darstellung über dem der normalen SOM liegt. Eigentlich läge im ersten Versuch die SOM-Darstellung mit farbigen Hexagons vorn. Im zweiten Versuch werden beide Ergebnisse durch zwei Ausreißer beeinflusst. An der Rangfolge ändert dies aber nichts. Einzig die Fehleranzahl wäre bei beiden Darstellungen



## Beste Visualisierung Datensatz COIL-100 und Fensterdatensatz

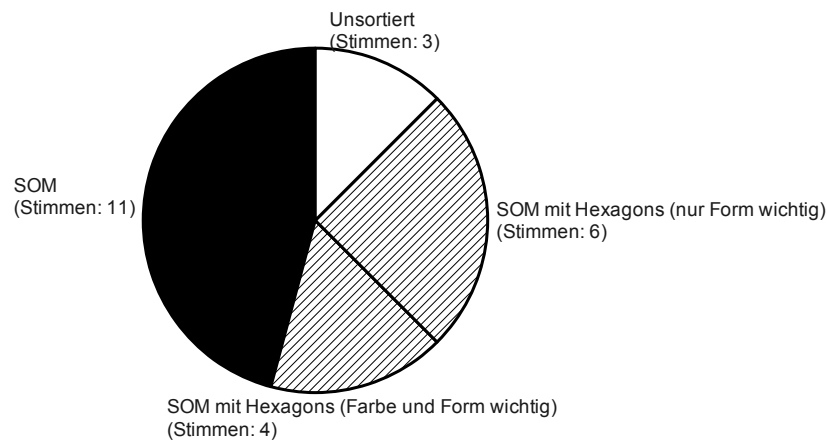


Abbildung 4.14: Abstimmungsergebnis der der Befragung zur besten Darstellung für das Labeln des COIL-100- und Fensterdatensatzes. Die Stimmen wurden nach jedem Datensatz abgegeben. Es konnte eine der verwendeten Darstellungen, als am besten geeignete, ausgewählt werden. Wenn die SOM-Darstellung mit farbigen Hexagons gewählt wurde, sollte der Tester außerdem entscheiden was bei dieser Darstellung am wichtigsten war: die Farbe, die Form oder beides. Die Farbe allein hat keine Stimmen bekommen, deswegen erscheint sie nicht im Diagramm.

eigentlich deutlich niedriger. Im dritten Versuch wird der Unterschied zwischen normaler SOM-Darstellung und Hexagon-SOM-Darstellung durch einen kleineren Ausreißer auf Seiten der zweiten Darstellung vergrößert. Eigentlich lägen die Ergebnisse nahezu auf gleicher Höhe.

Unter Einbeziehung der Ausreißer wird der Unterschied zwischen den beiden Darstellungen weniger klar. Die normale SOM-Darstellung liegt zwar im zweiten Versuch deutlich und im dritten Versuch leicht vorn, jedoch hatte die SOM-Darstellung mit farbigen Hexagons im ersten Versuch weniger Fehler. Da der zweite Versuch jedoch aus mehr Testergebnissen besteht, kann die normale SOM-Darstellung als die weniger Fehleranfällige von beiden gesehen werden.

Anschließend an das Labeln des COIL-100 und des Fensterdatensatzes wurden die Testpersonen befragt, welche Darstellung die beste für den jeweiligen Datensatz ist. Dabei standen die 3 gerade bearbeiteten Darstellungen zur Auswahl. Die gesammelten Stimmen sind in Diagramm 4.14 auf Seite 33 zu sehen. Die normale SOM-Darstellung liegt mit 11 von 24 Stimmen knapp vor der SOM Darstellung mit farbigen Hexagons,

welche 10 Stimmen erhielt. Diese 10 Stimmen unterteilen sich in drei Gruppen. Wenn eine Testperson die Hexagon-SOM-Darstellung als beste Darstellung wählte, sollte sie angeben was an den Hexagons am wichtigsten war. Die Farbe, die Form oder beides zusammen. Die Farbe allein erhielt keine Stimme, darum ist sie auch nicht im Diagramm vertreten. 6 der 10 Tester empfanden nur die Form der Hexagons hilfreich bzw. die Struktur, welche sie erzeugen. 4 Tester empfanden die Hexagonform und auch die Farbe als hilfreich. Vor allem die Farbwahl der Hexagons wurde mehrfach als störend benannt. Die Farben schienen teilweise zu kräftig zu sein und lenkten von den Bildern ab. Da das Farbspektrum viel Rot enthielt, waren außerdem die Auswahlwerkzeuge, welche ebenfalls Rot sind, schlecht zu sehen. Aus diesem Grund wurde, wie in Abschnitt 3.3.3 auf Seite 11 beschrieben, ein anderes Farbspektrum entworfen. Ein weiterer Kritikpunkt an der Hexagon-SOM-Darstellung war, dass die Farben zu selten einen Vorteil beim Labeln brachten. Der Wert, aus welchem die Farbe berechnet wird, bezieht sich auf alle umliegenden Knoten. Dadurch existieren nur wenige Knoten mit sehr eindeutigen Farben. Denn sie müssen dazu entweder inmitten einer Gruppe ähnlicher Knoten liegen, oder an der Grenze zwischen zwei unterschiedlichen Gruppen. Die übrigen Knotenpunkte erhalten oft Mischfarben, welche dem Benutzer wenig Rückschlüsse erlauben.

Insgesamt scheint die normale SOM-Darstellung die beste der drei Darstellungen zu sein. Jedoch ohne großen Vorsprung. Aus diesem Grund kann in der Anwendung zwischen allen drei Darstellungen gewechselt werden. So kann der Benutzer selbst bestimmen, mit welcher Darstellung er arbeiten möchte. Die SOM-Darstellung mit verschobenen Knoten und die SOM-Darstellung mit farbigen Hexagons bieten beide noch Potential zur Weiterentwicklung bzw. Verbesserung. Möglichkeiten dafür wurden bereits genannt und werden im Ausblick weiter ausgeführt.

### 4.2.3 Bewertung der Bildstatusfunktion

Die Möglichkeit den Status jedes Bildes zwischen „Aktiv“ und „Inaktiv“ zu wechseln soll dem Benutzer das Labeln vereinfachen. Da gelabelte Bilder automatisch deaktiviert werden, können die übrigen Bilder einfacher ausgewählt werden. Außerdem erhöht sich die Übersichtlichkeit, da deaktivierte Bilder weniger auffallen als aktivierte. Ob diese Annahmen auch von den Testern so empfunden wurden sollte mit der Evaluation ebenfalls herausgefunden werden. Der dazu verwendete kleinere COIL-Datensatz bestand aus 5 Objekten und insgesamt 432 Bildern.

Die zum Labeln benötigte Zeit wurde gemessen und in Diagramm 4.15 auf Seite 35 eingetragen. Dargestellt sind Mittelwerte aus jeweils 6 Ergebnissen. Im ersten Versuch ist nahezu kein Unterschied zwischen ein- und ausgeschalteter Bildstatusfunktion zu erkennen. Im zweiten Versuch wird der Unterschied jedoch deutlicher. Mit Bildstatusfunktion benötigten die Testpersonen weniger Zeit als ohne. Es ist auch erkennbar, dass die Geschwindigkeitssteigerung vom 1. zum 2. Versuch unterschiedlich groß ist. Wurde im



Abbildung 4.15: Vergleich der Labelzeit mit und ohne Bildstatusfunktion. Bei eingeschalteter Bildstatusfunktion konnte der Bildstatus zwischen „Aktiv“ und „Inaktiv“ gewechselt werden. Ungelabelte Bilder sind normalerweise „Aktiv“. Gelabelte Bilder werden automatisch auf „Inaktiv“ gesetzt, können aber auch wieder aktiviert werden. Die Zeiten wurden mit dem kleineren COIL-100-Datensatz, bestehend aus 5 Objekten und 432 Bildern, gesammelt. Die dargestellten Werte sind Mittelwerte aus jeweils 6 Testergebnissen.

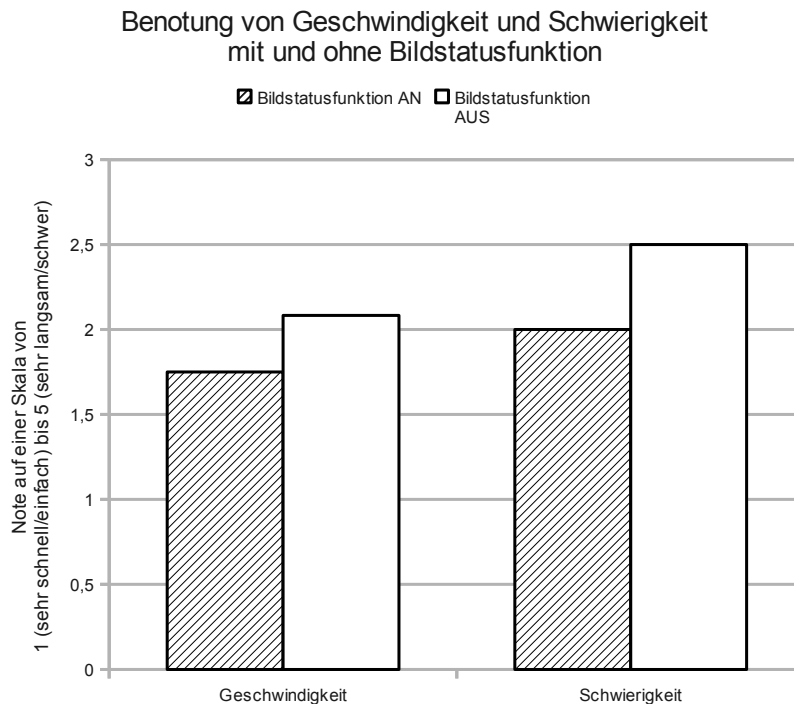


Abbildung 4.16: In der Befragung zu jeder Darstellung sollten die Tester ihren Eindruck zu Geschwindigkeit und Schwierigkeit, mit und ohne Bildstatusfunktion, angeben. Dazu stand eine fünfteilige Skala von „sehr schnell bzw. einfach“ bis „sehr langsam bzw. schwer“ zur Verfügung. Die dargestellten Werte sind Mittelwerte aus je 12 Ergebnissen.

ersten Versuch mit Bildstatusfunktion gearbeitet, so konnte die Zeit im zweiten Versuch nur geringfügig verbessert werden. Und das, obwohl die Anordnung der Bilder im zweiten Versuch exakt gleich war. Wurde im zweiten Versuch hingegen mit Bildstatusfunktion gearbeitet, ist eine größere Geschwindigkeitssteigerung möglich gewesen.

Nach dem Labeln wurden die Tester zu der gerade bearbeiteten Darstellung befragt. Sie sollten einschätzen wie schnell sie gelabelt haben und wie einfach das Labeln war. Dazu stand eine fünfteilige Skala von „sehr schnell bzw. einfach“ bis „sehr langsam bzw. schwer“ zur Verfügung. In Diagramm 4.16 auf Seite 36 sind die Mittelwerte der Ergebnisse abgebildet. Jeder Wert wurde aus 12 Antworten gebildet. Man kann erkennen, dass die eingeschaltete Bildstatusfunktion als schneller und leichter empfunden wurde. Der geringe Vorsprung der eingeschalteten Bildstatusfunktion ist wahrscheinlich durch die geringe Größe des Datensatzes bedingt. In den allgemeinen Befragungen wurde die Bildstatusfunktion als positiv eingeschätzt.

#### 4.2.4 Weitere Auswertungen

In der Abschließenden Befragung konnten die Testpersonen ihre Meinung zur Kompassfunktion äußern. Es sollte festgestellt werden, ob diese Funktion hilfreich ist oder nicht.

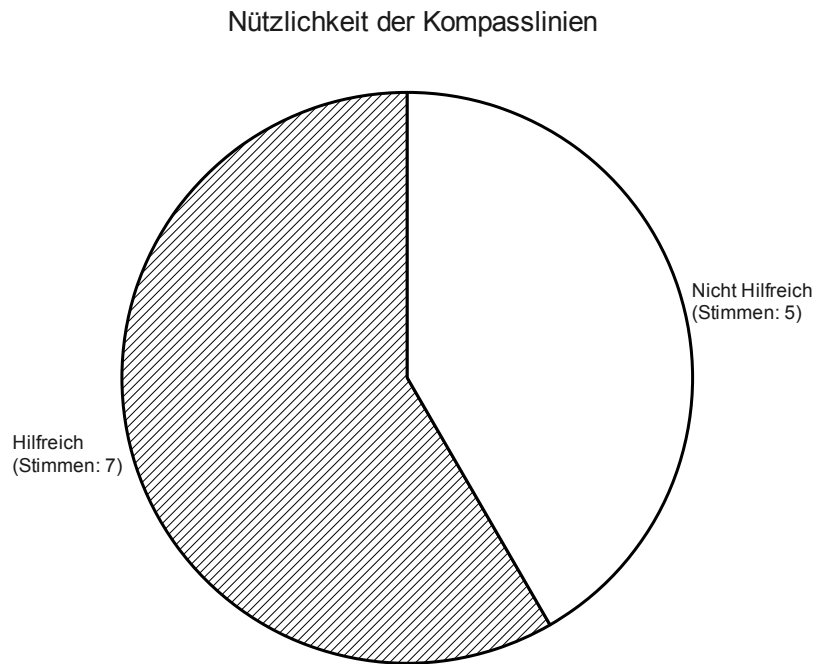


Abbildung 4.17: Ergebnisse der abschließenden Befragung zur Kompassfunktion. Keiner der Tester empfand die Kompasslinien als störend. Die Antworten wurden in die Gruppen „hilfreich“ und „nicht hilfreich“ unterteilt. Die Summe der Antworten in beiden Gruppen ist im Diagramm dargestellt.

Wichtig herauszufinden war außerdem, ob die Kompasslinien stören. Denn dann hätten sie in der normalen Anwendungsversion abschaltbar sein müssen. Dies war jedoch nicht der Fall. Die Antworten wurden daraufhin in die Gruppen „hilfreich“ und „nicht hilfreich“ aufgeteilt. Das Ergebnis ist in Diagramm 4.17 auf Seite 37 zu sehen. Eine knappe Mehrheit empfindet die Kompassfunktion als hilfreich. 5 von 12 Testern hat sie nicht geholfen. Aus Beobachtungen während des Labelns lässt sich eine Vermutung aufstellen. Es existieren scheinbar zwei verschiedene Vorgehensweisen beim Labeln. Eine Gruppe der Testpersonen verwendete vorrangig die Zoomfunktion, und sehr wenig die Verschiebefunktion. Die Tester zoomten zu jedem Knoten hinein, überprüften die zugehörigen Bilder und labelten diese. Danach zoomten sie wieder raus und wiederholten den Vorgang für den nächsten Knoten. Die andere Gruppe verwendete hauptsächlich die Verschiebefunktion und zoomte nur wenig. Die Tester dieser Gruppe zoomten hinein, bis zu einer Stufe bei der sie genug erkennen konnten. Danach bewegten sie sich fast nur in dieser Zoomstufe und labelten die Bilder. Dieser zweiten Gruppe könnte die Kompassfunktion nützlicher sein, als der ersten. Um diese Vermutung zu überprüfen wurden in Diagramm 4.18 auf Seite 38 die Anzahlen der Zoom- und Verschiebevorgänge gegenübergestellt. An der Y-Achse ist vermerkt, ob die Testperson den Kompass hilfreich oder nicht hilfreich fand. Nach der Vermutung müssten Tester mit einer hohen Summe Zoomvorgänge den Kompass nicht als hilfreich empfunden haben. Umgekehrt sollten Tester mit einer hohen Summe Verschiebevorgänge den Kompass hilfreich gefunden haben. Dies trifft allerdings

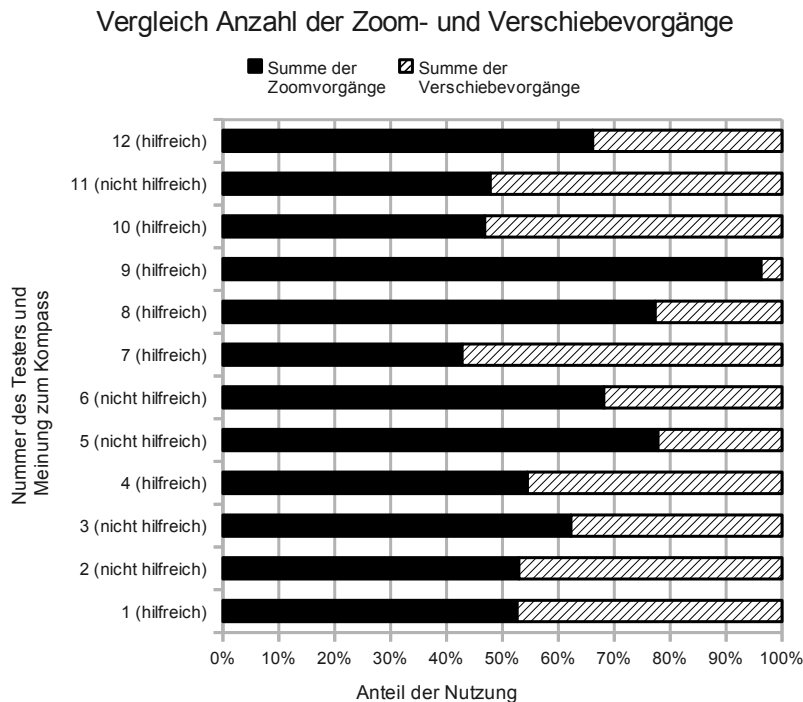


Abbildung 4.18: Zu sehen ist eine Gegenüberstellung der Summe der Zoomvorgänge und der Summe der Verschiebevorgänge jedes Testers. Es bestand die Annahme, dass Testpersonen, welche viel mit Zoom arbeiten die Kompassfunktion nicht hilfreich finden. Umgekehrt könnten Personen, welche die Ansicht viel verschieben, die Kompassfunktion hilfreich finden. Dazu ist bei der Nummer jeder Testperson vermerkt, ob sie den Kompass als „hilfreich“ oder „nicht hilfreich“ bewertet hat.

nicht zu. Nur 4 von 12 Testern (7,6,5,3) passen zur Annahme. Bei Tester 8 und 9 ist es sogar komplett umgekehrt. Die Annahme muss daher erweitert werden. Es kann sein, dass diese zwei Vorgehensweisen existieren. Jedoch scheint die Kompassfunktion auch nützlich zu sein, wenn man mit der anderen Vorgehensweise labelt. Außerdem verwendet ein Großteil der Benutzer sowohl Zoom als auch Verschieben etwa gleich viel.

Zusätzlich zur Kompassfunktion sollten die Testpersonen auch ihre Meinung zu den beiden Auswahlwerkzeugen „Pfad“ und „Rechteck“ äußern. Es bestand die Annahme, dass das Pfadwerkzeug besser geeignet ist, um in der Struktur der SOM eine Auswahl zu ziehen. Dies scheint jedoch nicht zuzutreffen. In Diagramm 4.19 auf Seite 39 erhält das Rechteckauswahlwerkzeug 8 von 12 Stimmen. Zwei Testpersonen empfanden das Pfadwerkzeug als besser, zwei andere meinten beide Werkzeuge hätten ihren Einsatzzweck und sind hilfreich. Oft wurde die Wahl des Rechteckwerkzeuges damit begründet, dass dieses gewohnter zu benutzen ist. In vielen Programmen werden Rechtecke zur Auswahl verwendet. Die Pfadauswahl hingegen war schwieriger zu ziehen. Dies könnte daher kommen, dass bei der Rechteckauswahl eigentlich nur eine Diagonale gezogen wird. Zwischen deren Start- und Endpunkt spannt sich dann ein Rechteck auf. Beim Pfadauswahlwerkzeug muss der gesamte Pfad mit der Maus nachgezogen

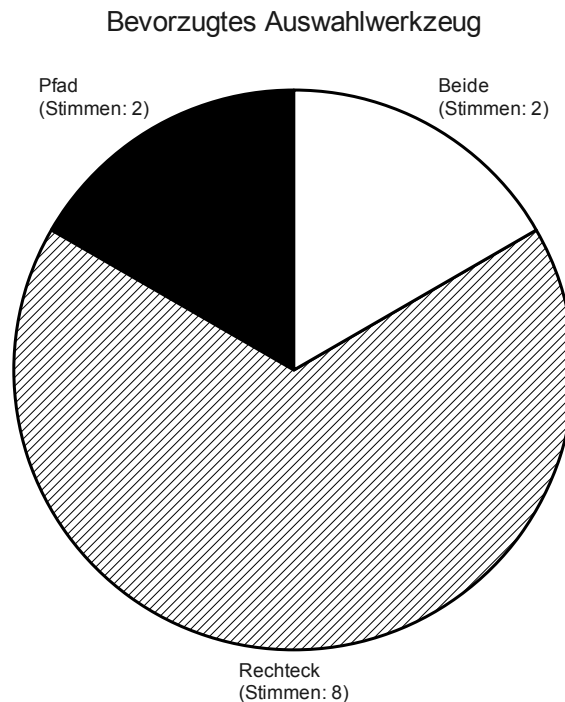


Abbildung 4.19: Ergebnis der Befragung nach dem favorisierten Auswahlwerkzeug. Die Antworten beziehen sich auf alle Datensätze und alle Darstellungen. Testpersonen, welche „beide“ wählten, begründeten dies mit den unterschiedlichen Einsatzmöglichkeiten der beiden Auswahlwerkzeuge.

werden.

Einige der aufgezeichneten Daten wurden in dieser Auswertung nicht verwendet. So zum Beispiel die Verwendungszeit des Rechteckwerkzeuges und des Pfadwerkzeuges. Die Auswertung dieser Zeiten erschien nicht mehr sinnvoll, nachdem die Ergebnisse der Befragung so eindeutig ausfielen. Der Vergleich der beiden Auswahlwerkzeuge entschied schließlich nur darüber, welches von beiden als Standard aktiviert ist. Genauso verhielt es sich bei den anderen nicht verwendeten Messdaten. Zu Beginn der Evaluation erschien es sinnvoll so viele Daten wie möglich aufzuzeichnen.

### 4.3 Anpassungen nach der Evaluation

Im Laufe der Evaluation und deren Auswertung ergaben sich mehrere Anregungen und Wünsche für Änderungen und neue Funktionen. Diese wurden gesammelt und Wünsche mit mehr als zwei Stimmen sind in Tabelle 4.3 auf Seite 40 aufgeführt. Im Anschluss an die Evaluation wurden dann einige dieser Wünsche implementiert.

Sehr deutlich ist der Wunsch nach anderen Farben in der Hexagon-SOM-Darstellung. Wie bereits unter 3.3.3 auf Seite 11 beschrieben, wurde darum ein anderes Farbspektrum

Vorschlag	Stimmen
Shortcuts zum schnelleren Vergeben von Labels	3
Anzahl der Bilder pro Knoten anzeigen	4
Bilder sollten sich nicht überlagern	5
Bildgröße beim Zoomen anpassen	7
Andere/keine Farben für Hexagons	11

Tabelle 4.3: Die 5 am häufigsten genannten Verbesserungsvorschläge. Während der Evaluation wurden mehrere Verbesserungsvorschläge von den Testern gemacht. Die Vorschläge mit mehr als 3 Nennungen sind in der Tabelle aufgeführt.

gewählt. Zusätzlich können die Hexagons auch komplett in einem Grauton dargestellt werden, ohne die U-Matrix Werte zu verwenden.

Einige Wünsche hätten zu große Änderungen an der Anwendung erfordert und konnten daher nicht umgesetzt werden. Dazu gehört der Wunsch nach Änderung der Bildgröße beim Zoom und der Wunsch nach einer Anordnung bei der sich keine Bilder überlagern. Auch der Wunsch nach einer Shortcutfunktion, bei der die Labels durch Drücken einer Taste auf der Tastatur zugewiesen werden können, war in der begrenzten Zeit nach der Evaluation nicht mehr zu realisieren.

Ein weiterer Wunsch war, dass auf den Repräsentanten der Knoten angezeigt wird, wie viele Bilder zu diesem Knoten gehören. Knoten mit nur einem Bild könnten sofort gelabelt werden, da der Repräsentant das einzige Bild des Knotens darstellt. Diese Funktion wurde implementiert und kann nun bei Bedarf aktiviert werden.

Zusätzlich wurden zwei kleinere Wünsche realisiert, welche leicht umzusetzen waren und hilfreich erschienen. Der eine Wunsch betraf das Selektieren der Bilder. Wenn man in einer hohen Zoomstufe alle Bilder eines Knotens auswählen wollte, so musste eine Auswahl um alle Bilder gezogen werden. Dies konnte, je nach Zoomstufe, eine sehr große Auswahlform erfordern. Um diese unnötig aufwendige Auswahl zu vermeiden, wurden die auf Knotenpunkten angezeigten Formen mit einer Knopffunktion versehen. Bei Klick auf die Form, werden alle Bilder des Knotens ausgewählt. Eine solche Knotenpunktform ist in der Mitte von Abbildung 4.8 auf Seite 26 zu sehen. Eine andere Funktion soll das Zoomen auf einen Knoten vereinfachen. Durch einen Doppelklick auf einen Repräsentanten wird ein Zoomvorgang gestartet. Dieser bewegt den zugehörigen Knoten schrittweise in die Mitte der Ansicht und zoomt gleichzeitig hinein bis zu einem bestimmten Zoomlevel. Dieser Vorgang ergibt eine Animation, welche Orientierungsverlust beim Benutzer vermeiden soll. Denn ein plötzlicher Wechsel von der Repräsentantenansicht zu einem einzelnen Knoten wäre zu irritierend.

Auf einige der Wünsche, welche nicht realisiert werden konnten, wird im Ausblick noch einmal eingegangen.



## 5 Zusammenfassung und Ausblick

Ziel dieser Arbeit war die Entwicklung einer Anwendung, mit welcher das Labeln von großen Bilddatensätzen vereinfacht wird. Die in Abschnitt 2.5 auf Seite 6 aufgestellten Anforderungen wurden alle erfüllt.

Nach der Entwicklung wurde die Evaluation zur Überprüfung der Vorteile durchgeführt. In dieser erhielten die entwickelten SOM-Darstellungen durchaus positive Ergebnisse, jedoch teilweise nicht so gravierend wie erwartet. Die Vorsortierung der Bilder wurde vor allem durch die Benutzer als bessere Darstellung empfunden, was ein sehr positives Zeichen ist. Aber sowohl die zum Labeln benötigte Zeit, als auch die Korrektheit der vergebenen Label ist noch verbesserungswürdig.

Besonders die Repräsentanten schienen in ihrer jetzigen Art wenig hilfreich. Anstatt nur ein Bild des Knotens anzuzeigen, sollten diese eine Art kombinierte Darstellung aller Bilder des Knotens sein. Wichtig für den Benutzer ist, dass erkennbar wird was für Bilder diesem Knoten angehören und ob es falsch zugeordnete Bilder gibt. Das Ziel der Repräsentanten sollte darum sein, dass nur bei fehlerhaft zugeordneten Bildern hineingezoomt werden muss und die meisten Knoten nach einem Blick auf den Repräsentanten richtig gelabelt werden können. Das Überprüfen der Knoten sollte daher auch weiter vereinfacht werden. In Abschnitt 4.3 auf Seite 39 sind bereits einige Funktionen aufgeführt, welche zu diesem Zweck nachträglich implementiert wurden. Eine weitere Idee wäre die Möglichkeit per Tastatur von Knotenpunkt zu Knotenpunkt zu wechseln. Mit den Pfeiltasten könnte eine Richtung vorgegeben werden, in welcher der nächstgelegene Knoten gesucht wird. Anschließend wird die Darstellung so verschoben, dass der Knoten im Zentrum der Anzeige liegt. Damit könnten übermäßig viele Verschiebevorgänge mit der Maus vermieden werden.

Einige Wünsche, welche während der Evaluation aufkamen konnten nicht mehr realisiert werden. Diese benötigen teilweise sehr große Änderungen an der Anwendung und könnten daher Thema einer nachfolgenden Arbeit sein. Einer dieser Wünsche, die Funktion, dass Bilder beim Zoomen vergrößert oder verkleinert werden, entspricht eigentlich nicht dem Konzept der Anwendung (Siehe Abschnitt 3.5 auf Seite 13). Interessant ist jedoch, wieso dieser Wunsch besteht. Auf Nachfrage meinten die Testpersonen, dass einige Bilder in der verwendeten Größe schlecht zu erkennen sind. Um diese Bilder genauer anzusehen, hätten sie gern die Funktion, dass diese beim Zoomen größer und kleiner werden. Für diese Funktionalität ist es aber nicht zwingend nötig die Größe aller Bilder zu ändern. Ebenso könnte eine Funktion entwickelt werden, mit welcher einzelne Bilder kurzzeitig vergrößert werden können. Dies wäre ähnlich zu einem Lupenwerkzeug.

Ein weiterer Wunsch war, dass die Bilder in der SOM-Darstellung mit verschobenen Kno-

tenpunkten einander nicht überlagern. Es wäre nützlich eine Anordnung zu entwickeln, bei der sich Bilder automatisch abstoßen. In so einer Anordnung könnten Knotenpunkte und Bilder durch Beziehungen, ähnlich wie Anziehungs- und Abstoßungskräfte, verbunden sein. Ebenso die einzelnen Bilder untereinander. Durch eine solche Anordnung würden sich Überlagerungen automatisch auflösen.

Insgesamt scheint die Vorsortierung von Bildern mithilfe einer self organizing map sehr sinnvoll. Durch die genannten Änderungen und weitere Entwicklungen kann diese große Vorteile beim Kategorisieren großer Datensätze bringen.

## Literaturverzeichnis

- [Heidemann et al., 2003] Gunther Heidemann, Axel Saalbach, Helge Ritter: Semi-automatic acquisition and labelling of image data using SOMs, ESANN'2003 proceedings, 2003
- [Bekel et al., 2005] Holger Bekel, Gunther Heidemann, Helge Ritter: Interactive Image Data Labeling Using Self-Organizing Maps in an Augmented Reality Scenario, Neural Networks Vol. 18(5/6), 2005
- [Kohonen, 1990] Teuvo Kohonen: The self organizing map, Proceedings of the IEEE, 1990
- [Kessler, 2010] Stefanie Wiltrud Kessler: Analysis and Visualization of Coreference Features, Diplomarbeit, Universität Stuttgart, Stuttgart 2010
- [Heer J. et al., 2005] Jeffrey Heer, Stuart K. Card, James Landay: Prefuse: A Toolkit for Interactive Information Visualization, 2005
- [Nene et al., 1996] Sameer A. Nene, Shree K. Nayar, Hiroshi Murase: Columbia Object Image Library: COIL-100, Technical Report No. CUCS- 006-96, Dept. Computer Science, Columbia University
- [1] Columbia Object Image Library (COIL-100):  
<http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php>
- [2] Universität Paderborn - TowerCam: <http://towercam.upb.de/parkplatz.html>



## Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Mittweida, 01. August 2010